BIOCHEMISTRY

Second Edition

SCHAUM'S OUTLINE OF

THEORY AND PROBLEMS

OF

BIOCHEMISTRY

Second Edition

•

PHILIP W. KUCHEL, Ph.D.
Coordinating Author

GREGORY B. RALSTON, Ph.D.
Coordinating Author

AUDREY M. BERSTEN, M.Sc.

SIMON B. EASTERBROOK-SMITH, Ph.D.

ALAN R. JONES, Ph.D.

M. DAN MONTAGUE, Ph.D.

MICHAEL B. SLAYTOR, Ph.D.

MICHAEL A. W. THOMAS, D.Phil.

R. GERARD WAKE, Ph.D.

With new material by:

DOUGLAS J. CHAPPELL, Ph.D.

RICHARD I. CHRISTOPHERSON, Ph.D.

ARTHUR D. CONIGRAVE, Ph.D.

IVAN G. DARVEY, Ph.D.

J. MITCHELL GUSS, Ph.D.

GLENN F. KING, Ph.D.

MICHAEL B. MORRIS, Ph.D.

SAMIR SAMMAN, Ph.D.

MARK T. SMITH, B.Sc.

EVE SZABADOS, Ph.D.

ANTHONY S. WEISS, Ph.D.

EMMA WHITELAW, D.Phil

Department of Biochemistry
The University of Sydney
Sydney, Australia

•

SCHAUM'S OUTLINE SERIES

McGRAW-HILL

New York San Francisco Washington, D.C. Auckland Bogotá Caracas Lisbon
London Madrid Mexico City Milan Montreal New Dehli
San Juan Singapore Sydney Tokyo Toronto

PHILIP W. KUCHEL, Ph.D., and **GREGORY B. RALSTON**, Ph.D., are Professors of Biochemistry at the University of Sydney, Australia. They coordinated the writing of this book, with contributions from seven other members of the teaching staff, and editorial assistance from many more, in the Department of Biochemistry at the University.

Schaum's Outline of Theory and Problems of
BIOCHEMISTRY

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 PRS PRS 9 0 2 1 0 9 8

ISBN 0-07-036149-5

Sponsoring Editor: Barbara Gilson
Production Supervisor: Sherri Souffrance
Editing Supervisor: Maureen B. Walker

Library of Congress Cataloging-in-Publication Data

Schaum's outline of theory and problems of biochemistry / Philip W.
Kuchel . . . [et al.].—2nd ed.
p. cm.—(Schaum's outline series)
Includes index.
ISBN 0-07-036149-5 (pbk.)
1. Biochemistry—Outlines, syllabi, etc. 2. Biochemistry—
—Examinations, questions, etc. 1. Kuchel, Philip W.
QP518.3.S3 1997 97-23525
572—dc21 CIP

McGraw-Hill
A Division of The McGraw-Hill Companies

Preface to the Second Edition

In the time since the first edition of the book, biochemistry has undergone great developments in some areas, particularly in molecular biology, signal transduction, and protein structure. Developments in these areas have tended to overshadow other, often more traditional, areas of biochemistry such as enzyme kinetics. This second edition has been prepared to take these changes in direction into account: to emphasize those areas that are rapidly developing and to bring them up to date. The preparation of the second edition also gave us the opportunity to adjust the balance of the book, and to ensure that the depth of treatment in all chapters is comparable and appropriate for our audiences.

The major developments in biochemistry over the last 10 years have been in the field of molecular biology, and the second edition reflects these changes with significant expansion of these areas. We are very grateful to Dr. Emma Whitelaw for her substantial efforts in revising Chapter 17. In addition, increased understanding of the dynamics of DNA structures, developments in recombinant DNA technology, and the polymerase chain reaction have been incorporated into the new edition, thanks to the efforts of Drs. Anthony Weiss and Doug Chappell. The section on proteins also has been heavily revised, by Drs. Glenn King, Mitchell Guss, and Michael Morris, reflecting significant growth in this area, with greater emphasis on protein folding. A number of diagrams have been redrawn to reflect our developing understanding, and we are grateful to Mr. Mark Smith and to Drs. Eve Szabados and Michael Morris for their art work.

The sections on lipid metabolism, membrane function, and signal transduction have been enlarged and enhanced, reflecting modern developments in these areas, through the efforts of Drs. Samir Samman and Arthur Conigrave. In the chapter on nitrogen metabolism, the section on nucleotides has been enlarged, and the coverage given to the metabolism of specific amino acids has been correspondingly reduced. For this we are grateful to Dr. Richard Christopherson.

In order to avoid excessive expansion of the text, the material on enzymology and enzyme kinetics has been refocused and consolidated, reflecting changes that have taken place in the teaching of these areas in most institutions. We are grateful to Dr. Ivan Darvey for his critical comments and helpful suggestions in this endeavor.

The style of presentation in the current edition continues that of the first edition, with liberal use of *didactic questions* that attempt to develop concepts from prior knowledge, and to promote probing of the gaps in that knowledge. Thus, the book has been prepared through the efforts of many participants who have contributed in their areas of specialization; we have been joined in this endeavor by several new contributors whose sections are listed above.

PHILIP W. KUCHEL
GREGORY B. RALSTON
Coordinating Authors

This page intentionally left blank

Preface to the First Edition

This book is the result of a cooperative writing effort of approximately half of the academic staff of the largest university department of biochemistry in Australia. We teach over 1,000 students in the Faculties of Medicine, Dentistry, Science, Pharmacy, Veterinary Science, and Engineering. So, for whom is this book intended and what is its purpose?

This book, as the title suggests, is an *Outline* of Biochemistry—principally mammalian biochemistry and not the full panoply of the subject. In other words, it is not an encyclopedia but, we hope, a guide to understanding for undergraduates up to the end of their B.Sc. or its equivalent.

Biochemistry has become the language of much of biology and medicine; its principles and experimental methods underpin all the basic biological sciences in fields as diverse as those mentioned in the faculty list above. Indeed, the boundaries between biochemistry and much of medicine have become decidedly blurred. Therefore, in this book, either implicitly through the solved problems and examples, or explicitly, we have attempted to expound *principles* of biochemistry. In one sense, this book is our definition of biochemistry; in a few words, we consider it to be the description, using *chemical* concepts, of the processes that take place in and by living organisms.

Of course, the chemical processes in cells occur not only in free solution but are associated with macromolecular structures. So inevitably, biochemistry must deal with the structure of tissues, cells, organelles, and of the individual molecules themselves. Consequently, this book begins with an overview of the main procedures for studying cells and their organelle constituents, with what the constituents are and, in general terms, what their biochemical functions are. The subsequent six chapters are far more chemical in perspective, dealing with the major classes of biochemical compounds. Then there are three chapters that consider enzymes and general principles of metabolic regulation; these are followed by the metabolic pathways that are the real soul of biochemistry.

It is worth making a few comments on the *style* of presenting the material in this book. First, we use so-called *didactic questions* that are indicated by the word *Question*; these introduce a new topic, the answers for which are not available from the preceding text. We feel that this approach embodies and emphasizes the inquiry in any research, including biochemistry: the answer to one question often immediately provokes another question. Secondly, as in other Schaum's Outlines, the basic material in the form of *general* facts is emphasized by what is, essentially, optional material in the form of *examples*. Some of these examples are written as questions; others are simple expositions on a particular subject that is a specific example of the general point just presented. Thirdly, the solved problems relate, according to their section headings, to the material in the main text. In virtually all cases, students should be able to solve these problems, at least to a reasonable depth, by using the material in this outline. Finally, the supplementary problems are usually questions that have a minor twist on those already considered in either

of the previous three cataegories; answers to these questions are provided at the end of the book.

While this book was written by academic staff, its production has also depended on the efforts of many other people, whom we thank sincerely. For typing and word processing, we thank Anna Dracopoulos, Bev Longhurst-Brown, Debbie Manning, Hilary McDermott, Elisabeth Sutherland, Gail Turner, and Mary Walsh and for editorial assistance, Merilyn Kuchel. For critical evaluation of the manuscript, we thank Dr. Ivan Darvey and many students, but especially Tiina Iismaa, Glenn King, Kiaran Kirk, Michael Morris, Julia Raftos, and David Thorburn. Dr. Arnold Hunt helped in the early stages of preparing the text. We mourn the sad loss of Dr. Reg O'Brien, who died when this project was in its infancy. We hope, given his high standards in preparing the written and spoken word, that he would have approved of the final form of the book. Finally, we thank Elizabeth Zayatz and Marthe Grice of McGraw-Hill; Elizabeth for raising the idea of the book in the first place, and both of them for their enormous efforts to satisfy our publication requirements.

PHILIP W. KUCHEL
GREGORY B. RALSTON
Coordinating Authors

Contents

Chapter 1 CELL ULTRASTRUCTURE **1**

1.1 Introduction 1
1.2 Methods of Studying the Structure and Function of Cells 1
1.3 Subcellular Organelles 7
1.4 Cell Types 15
1.5 The Structural Hierarchy in Cells 17

Chapter 2 CARBOHYDRATES **25**

2.1 Introduction and Definitions 25
2.2 Glyceraldehyde 26
2.3 Simple Aldoses 27
2.4 Simple Ketoses 30
2.5 The Structure of D-Glucose 32
2.6 The Conformation of Glucose 35
2.7 Monosaccharides Other Than Glucose 38
2.8 The Glycosidic Bond 42
2.9 Polysaccharides 46

Chapter 3 AMINO ACIDS AND PEPTIDES **53**

3.1 Amino Acids 53
3.2 Acid-Base Behavior of Amino Acids 56
3.3 Amino Acid Analysis 65
3.4 The Peptide Bond 66
3.5 Reactions of Cysteine 68

Chapter 4 PROTEINS **76**

4.1 Introduction 76
4.2 Purification and Characterization of Proteins 76
4.3 Protein Folding 84
4.4 Protein Structure 87
4.5 Sequence Homology and Protein Evolution 97
4.6 Methods for Protein Structure Determination 99

Chapter 5 PROTEINS: SUPRAMOLECULAR STRUCTURE **108**

5.1 Introduction 108
5.2 Assembly of Supramolecular Structures 108
5.3 Protein Self-Association 111
5.4 Hemoglobin 117

5.5 The Extracellular Matrix 121
5.6 The Cytoskeleton 130

Chapter 6 LIPIDS, MEMBRANES, TRANSPORT, AND SIGNALING 153

6.1 Introduction 153
6.2 Classes of Lipids 154
6.3 Fatty Acids 155
6.4 Glycerolipids 157
6.5 Sphingolipids 161
6.6 Lipids Derived from Isoprene (Terpenes) 162
6.7 Behavior of Lipids in Water 165
6.8 Bile Acids and Bile Salts 168
6.9 Plasma Lipoproteins 169
6.10 Vesicles 170
6.11 Membranes 171
6.12 Transport 176
6.13 Molecular Mechanisms of Transport Across Membranes 182
6.14 Signaling 185

Chapter 7 NUCLEIC ACIDS 198

7.1 Introduction 198
7.2 Nucleic Acids and Their Chemical Constituents 198
7.3 Nucleosides 201
7.4 Nucleotides 202
7.5 Polynucleotides 204
7.6 Structure of DNA 206
7.7 Denaturation of DNA 212
7.8 Size, Organization, and Topology of DNA 215
7.9 Structure and Types of RNA 218
7.10 Nucleases 219

Chapter 8 ENZYME CATALYSIS 228

8.1 Basic Concepts 228
8.2 Classification of Enzymes 229
8.3 Modes of Enhancement of Rates of Bond Cleavage 230
8.4 Rate Enhancement and Activation Energy 237
8.5 Site-Directed Mutagenesis 238

Chapter 9 ENZYME KINETICS 251

9.1 Introduction and Definitions 251
9.2 Dependence of Enzyme Reaction Rate on Substrate Concentration 252
9.3 Graphical Evaluation of K_m and V_{max} 253
9.4 Enzyme Inhibition—Definitions 254
9.5 Enzyme Inhibition—Equations 255
9.6 Mechanistic Basis of the Michaelis-Menten Equation 255
9.7 Derivation of Complicated Steady-State Equations 257
9.8 Multireactant Enzymes 259

9.9 pH Effects on Enzyme Reaction Rates 261
9.10 Mechanisms of Enzyme Inhibition 263
9.11 Regulatory Enzymes 265

Chapter *10* METABOLISM: UNDERLYING THEORETICAL PRINCIPLES 290

10.1 Introduction 290
10.2 Thermodynamics 290
10.3 Redox Reactions 295
10.4 ATP and Its Role in Bioenergetics 298
10.5 Control Points in Metabolic Pathways 299
10.6 Amplification of Control Signals 301
10.7 Intracellular Compartmentation and Metabolism 303

Chapter *11* CARBOHYDRATE METABOLISM 311

11.1 Glycolysis 311
11.2 The Fate of Pyruvate 319
11.3 Gluconeogenesis 323
11.4 The Cori Cycle 326
11.5 Glycogen Metabolism 327
11.6 The Entry of Other Carbohydrates into Glycolysis 330
11.7 Regeneration of Cytoplasmic NAD^+ Levels 332
11.8 Control of Glycolysis 334
11.9 Effects of Hormones on Glycolysis 336
11.10 The Pentose Phosphate Pathway 339

Chapter *12* THE CITRIC ACID CYCLE 345

12.1 Introduction 345
12.2 Reactions of the Citric Acid Cycle 346
12.3 The Energetics of the Citric Acid Cycle 349
12.4 Regulation of the Citric Acid Cycle 350
12.5 The Pyruvate Dehydrogenase Complex 352
12.6 Pyruvate Carboxylase 353
12.7 The Amphibolic Nature of the Citric Acid Cycle 354
12.8 The Glyoxylate Cycle 355

Chapter *13* LIPID METABOLISM 362

13.1 Introduction 362
13.2 Lipid Digestion 362
13.3 Lipoprotein Metabolism 364
13.4 Mobilization of Depot Lipid 368
13.5 Oxidation of Fatty Acids 368
13.6 The Fate of Acetyl-CoA from Fatty Acids: Ketogenesis 370
13.7 Lipogenesis 374
13.8 Synthesis of Phospholipids and Sphingolipids 379
13.9 Prostaglandins 383
13.10 Metabolism of Cholesterol 387
13.11 Regulation of Lipid Metabolism 392

Chapter *14* OXIDATIVE PHOSPHORYLATION 402

14.1 Introduction 402
14.2 Components of the Electron-Transport Chain 402
14.3 Organization of the Electron-Transport Chain 405
14.4 Coupling of Electron Transport and ATP Synthesis 407
14.5 The Ratio of Protons Extruded from the Mitochondrion to Electrons Transferred to Oxygen 408
14.6 Mechanistic Models of Proton Translocation 409
14.7 ATP Synthase 412
14.8 The Mechanism of ATP Synthesis 412
14.9 Transport of Adenine Nucleotides to and from Mitochondria 414

Chapter *15* NITROGEN METABOLISM 419

15.1 Synthesis and Dietary Sources of Amino Acids 419
15.2 Digestion of Proteins 426
15.3 Dynamics of Amino Acid Metabolism 431
15.4 Amino Acid Catabolism 432
15.5 Disposal of Excess Nitrogen 434
15.6 Pyrimidine and Purine Metabolism 437
15.7 Metabolism of C_1 Compounds 447
15.8 Porphyrin Metabolism 451

Chapter *16* REPLICATION AND MAINTENANCE OF THE GENETIC MATERIAL 458

16.1 Introduction 458
16.2 Semiconservative Replication of DNA 458
16.3 Topology of DNA Replication 459
16.4 Control of DNA Replication 462
16.5 Enzymology of DNA Replication in Bacteria 464
16.6 Molecular Events in the Initiation of Replication in Bacteria 469
16.7 Termination of Chromosome Replication in Bacteria 470
16.8 Initiation, Elongation, and Termination of Replication in Eukaryotes 472
16.9 Inhibitors of DNA Replication 473
16.10 Repair of DNA Damage 475
16.11 Recombinant DNA and Isolation of Genes 476
16.12 The Polymerase Chain Reaction 477

Chapter *17* GENE EXPRESSION AND PROTEIN SYNTHESIS 489

17.1 Introduction 489
17.2 The Genetic Code 489
17.3 DNA Transcription in Bacteria 491
17.4 DNA Transcription in Eukaryotes 494
17.5 Transcription Factors 494
17.6 Processing the RNA Transcript 497
17.7 Organization of the Genome 498
17.8 Inhibitors of Transcription 499
17.9 The mRNA Translation Machinery 500
17.10 RNA Translation in Bacteria 503
17.11 RNA Translation in Eukaryotes 505

17.12 Posttranslational Modification of Proteins 505
17.13 Inhibitors of Translation 506
17.14 Control of Gene Expression 508

ANSWERS TO SUPPLEMENTARY PROBLEMS **519**

INDEX **539**

This page intentionally left blank

Chapter 1

Cell Ultrastructure

1.1 INTRODUCTION

Question: What are the basic units of life?

All animals, plants, and microorganisms are composed of small units known as *cells*. Cells range in volume from a few attoliters among bacteria to milliliters for the giant nerve cells of squid; typical cells in mammals have diameters of 10 to 100 μm and are thus often smaller than the smallest visible particle. They are generally flexible structures with a delimiting membrane that is in a dynamic, undulating state. Different animal and plant tissues contain different types of cells, which are distinguished not only by their different structure but by their different metabolic activities.

EXAMPLE 1.1

Antonie van Leeuwenhoek (1632–1723), draper of Delft in Holland, ground his own lenses and made simple microscopes that gave magnifications of ~×200. On October 9, 1676, he sent a 17½-page letter to the Royal Society of London, in which he described *animalcules* in various water samples. These small organisms included what are today known as protozoans and *bacteria*; thus Leeuwenhoek is credited with the first observation of bacteria. Later work of his included the identification of spermatozoa and red blood cells from many species.

The development of a *stem cell* into cells with specialized function is called the process of *differentiation*. This takes place most dramatically in the development of a fetus, from the single cell formed by the fusion of one *spermatozoon* and one *ovum* to a vast array of different tissues.

Cells appear to be able to recognize cells of like kind, and thus to unite into coherent organs, principally because of specialized glycoproteins (Chap. 2) on the cell membranes.

1.2 METHODS OF STUDYING THE STRUCTURE AND FUNCTION OF CELLS

Light Microscopy

Many cells and, indeed, parts of cells (*organelles*) react strongly with colored dyes such that they can be easily distinguished in thinly cut sections of tissue by using light microscopy. Hundreds of different dyes with varying degrees of selectivity for tissue components are used for this type of work, which constitutes the basis of the scientific discipline *histology*.

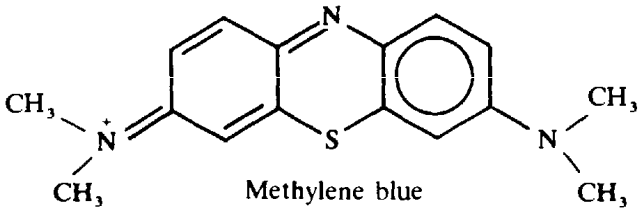
EXAMPLE 1.2

In the clinical biochemical assessment of patients, it is common practice to inspect a blood sample under the light microscope, with a view to determining the number and type of inflammatory white cells present. A thin film of blood is smeared on a glass slide, which is then placed in methanol to *fix* the cells; this process rigidifies the cells and preserves their shape. The cells are then dyed by the addition of a few drops each of two dye mixtures; the most commonly used ones are the *Romanowsky* dyes, named after their nineteenth-century discoverer. The commonly used hematological dyeing procedure is that developed by J.W. Field: A mixture of *azure I* and *methylene blue* is first applied to the cells, followed by *eosin*; all dyes are dissolved in a simple phosphate buffer. The treatment stains nuclei blue, cell cytoplasm pink, and some subcellular organelles either pink or blue. On the basis of different staining patterns, at least five different types of white cells can be identified. Furthermore, intracellular organisms such as the malarial parasite *Plasmodium* stain blue.

The exact chemical mechanisms of differential staining of tissues are poorly understood. This aspect of histology is therefore still empirical. However, certain features of the chemical structure of dyes allow some interpretation of how they achieve their selectivity. They tend to be multi-ring, heterocyclic, aromatic compounds, with the high degree of bond conjugation giving the bright colors. In many cases they were originally isolated from plants, and they have a net positive or negative charge.

EXAMPLE 1.3

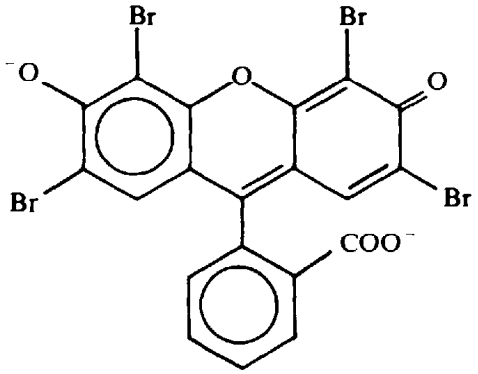
Methylene blue stains cellular nuclei blue.



Methylene blue

Mechanism of staining: The positive charge on the N of methylene blue interacts with the anionic oxygens in the phosphate esters of DNA and RNA (Chap. 7).

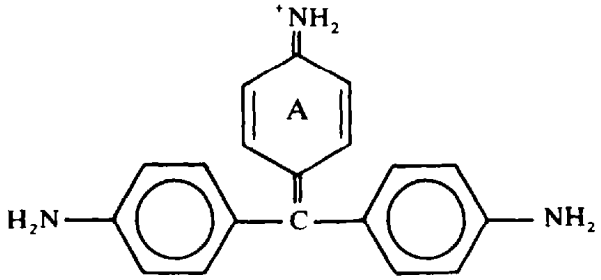
Eosin stains protein-rich regions of cells red.



Eosin

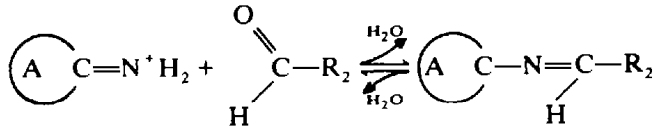
Mechanism of staining: Eosin is a *dianion* at pH 7, and so it binds electrostatically to protein groups, such as arginyls, histidyls, and lysyls, that have a positive charge at this pH. Thus, this dye highlights protein-rich areas of cells.

PAS (*periodic acid Schiff*) stain is used for the histological staining of carbohydrates; it is also used to stain glycoproteins (proteins that contain carbohydrates; Chap. 2) in electrophoretic gels (Chap. 4). The stain mixture contains *periodic acid* (HIO_4), a powerful oxidant, and the dye *basic fuchsin*:



Basic fuchsin

Mechanism of staining: Periodic acid opens the sugar rings at cis-diol bonds (i.e., the C-2—C-3 bond of glucose) to form two aldehyde groups and *iodate* (IO_3^-). Then the $=^+NH_2$ group of the dye reacts to form a so-called *Schiff base* bond with the aldehyde, thus linking the dye to the carbohydrate. The basic reaction is:



The conversion of ring A of basic fuchsin to an aromatic one, with a carbocation (positively charged carbon atom) at the central carbon, renders the compound pink.

Electron Microscopy

Image magnifications of thin tissue sections up to ×200,000 can be achieved using *electron microscopy*. The sample is placed in a high vacuum and exposed to a narrow beam of electrons that are differentially scattered by different parts of the section. Therefore, in staining the sample, differential electron density replaces the colored dyes used in light microscopy. A commonly used dye is *osmium tetroxide* (OsO_4), which binds to amino groups of proteins, leaving a black, electron-dense region.

EXAMPLE 1.4

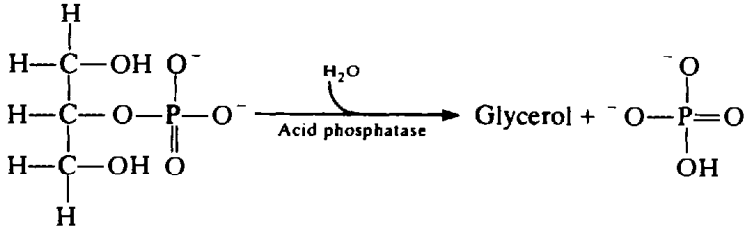
The wavelength of electromagnetic radiation (light) limits the resolution attainable in microscopy. The *resolution* of a device is defined as the smallest gap that can be perceived between two objects when viewed with the device; resolution is approximately half the wavelength of the electromagnetic radiation used. Electrons accelerated to high velocities by an electrical potential of ~100,000 V have electromagnetic wave properties as well, with a wavelength of 0.004 nm; thus a resolution of about 0.002 nm is theoretically attainable with electron microscopy. This, at least in principle, enables the distinction of certain features even on protein molecules, since the diameter of many globular proteins, e.g., hemoglobin, is greater than 3 nm; in practice, however, such resolution is not usually attained.

Histochemistry and Cytochemistry

Histochemistry deals with whole tissues, and *cytochemistry* with individual cells. The techniques of these disciplines give a means for locating specific compounds or enzymes in tissues and cells. A tissue slice is incubated with the substrate of an enzyme of interest, and the product of this reaction is caused to react with a second, pigmented compound that is also present in the incubation mixture. If the samples are adequately fixed before incubation and the fixing process does not damage the enzyme, the procedure will highlight, in a thin section of tissue under the microscope, those cells which contain the enzyme or, at higher resolution, the subcellular organelles which contain it.

EXAMPLE 1.5

The enzyme *acid phosphatase* is located in the lysosomes (Sec. 1.3) of many cells, including those of the liver. The enzyme catalyzes the hydrolytic release of phosphate groups from various phosphate esters including the following:



2-Phosphoglycerol Phosphate

In the *Gomori* procedure, tissue samples are incubated for ~30 min at 37°C in a suitable buffer that contains 2-phosphoglycerol. The sample is then washed free of the phosphate ester and placed in a buffer that contains lead nitrate. The 2-phosphoglycerol freely permeates lysosomal membranes, but the more highly charged phosphate does not, so that any of the phosphate released inside the lysosomes by phosphatase remains there. As the Pb^{2+} ions penetrate the lysosomes, they precipitate as lead phosphate. These regions of precipitation appear as dark spots in either an electron or a light micrograph.

Autoradiography

Autoradiography is a technique for locating radioactive compounds within cells; it can be conducted with light or electron microscopy. Living cells are first exposed to the *radioactive precursor* of some intracellular component. The labeled precursor is a compound with one or more hydrogen (^1H) atoms replaced by the *radioisotope tritium* (^3H); e.g., [^3H]thymidine is a labeled precursor of DNA, and [^3H]uridine is a labeled precursor of RNA (Chap. 7). Various tritiated amino acids are also available. The labeled precursors enter the cells and are incorporated into the appropriate macromolecules. The cells are then fixed, and the samples are embedded in a resin or wax and then sectioned into thin slices.

The radioactivity is detected by applying (in a darkroom) a photographic silver halide emulsion to the surface of the section. After the emulsion dries, the preparations are stored in a light-free box to permit the radioactive decay to expose the overlying emulsion. The length of exposure depends on the amount of radioactivity in the sample but is typically several days to a few weeks for light microscopy and up to several months for electron microscopy. The long exposure time in electron microscopy is necessary because of the *very* thin sections ($<1\ \mu m$) and thus the minute amounts of radioactivity present in the tiny samples. The preparations are developed and fixed as in conventional photography. Thus, the silver grains overlie regions of the cell that contain radioactive molecules; the grains appear as tiny black dots in light micrographs and as twisted black threads in electron micrographs. Note that this whole procedure works only if the precursor molecule can traverse the cell membrane and if the cells are in a phase of their life cycle that involves incorporation of the compound into macromolecules.

EXAMPLE 1.6

The sequence of events involved in the synthesis and transport of *secretory proteins* from glands can be followed using autoradiography. For example, rats were injected with [^3H]leucine, and at intervals thereafter they were sacrificed and autoradiographs of their *prostate glands* prepared. In electron micrographs of the sample obtained 4 min after the injection, silver grains appeared overlying the *rough endoplasmic reticulum* (RER) of the cells, indicating that [^3H]leucine had been incorporated from the blood into protein by the *ribosomes* attached to the RER. By 30 min the grains were overlying the Golgi apparatus and secretory vacuoles, reflecting intracellular transport of labeled secretory proteins from the RER to those organelles. At later times after the injection radioactive proteins were released from the cells, as evidenced by the presence of silver grains over the glandular lumens.

Ultracentrifugation

The *biochemical* roles of subcellular organelles could not be studied properly until the organelles had been separated by fractionation of the cells. George Palade and his colleagues, in the late 1940s, showed that *homogenates* of rat liver could be separated into several fractions using *differential centrifugation*. This procedure relies on the different velocities of sedimentation of various organelles of different shape, size, and density through a solution. A typical experiment is outlined in Example 1.7.

EXAMPLE 1.7

Liver is suspended in 0.25 *M* sucrose and then disrupted using a rotating, close-fitting Teflon plunger in a glass barrel (known as a *Potter-Elvehjem homogenizer*). Care is taken not to destroy the organelles by excessive homogenization. The sample is then spun in a centrifuge (see Fig. 1-1). The nuclei tend to be the first to sediment to the bottom of the sample tube at forces as low as 1,000 *g* for ~15 min in a tube 7 cm long.

High-speed centrifugation, such as 10,000 *g* for 20 min, yields a pellet composed mostly of mitochondria, but contaminated with lysosomes. Further centrifugation at 100,000 *g* for 1 h yields a pellet of ribosomes and so-called microsomes that contain endoplasmic reticulum. The soluble protein and other solutes remain in the *supernatant* (overlying solution) from this step.

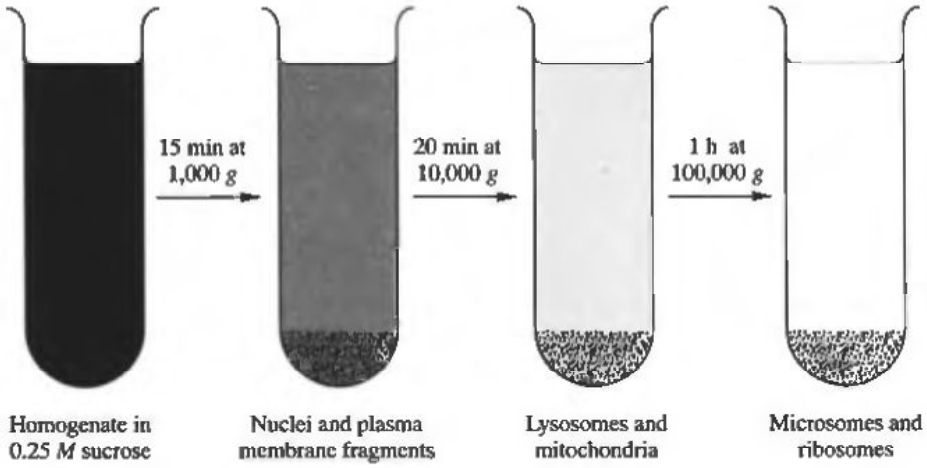


Fig. 1-1 Separation of subcellular organelles by differential centrifugation of cell homogenates.

Density-gradient centrifugation (also called *isopycnic centrifugation*) can also be used to separate the different organelles (Fig. 1-2). The homogenate is layered onto a discontinuous or continuous concentration gradient of sucrose solution, and centrifugation continues until the subcellular particles achieve density equilibrium with their surrounding solution.

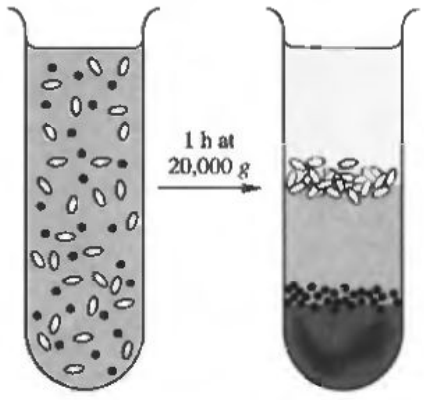


Fig. 1-2 Isopycnic centrifugation of organelles. The shading indicates increasing solution density.

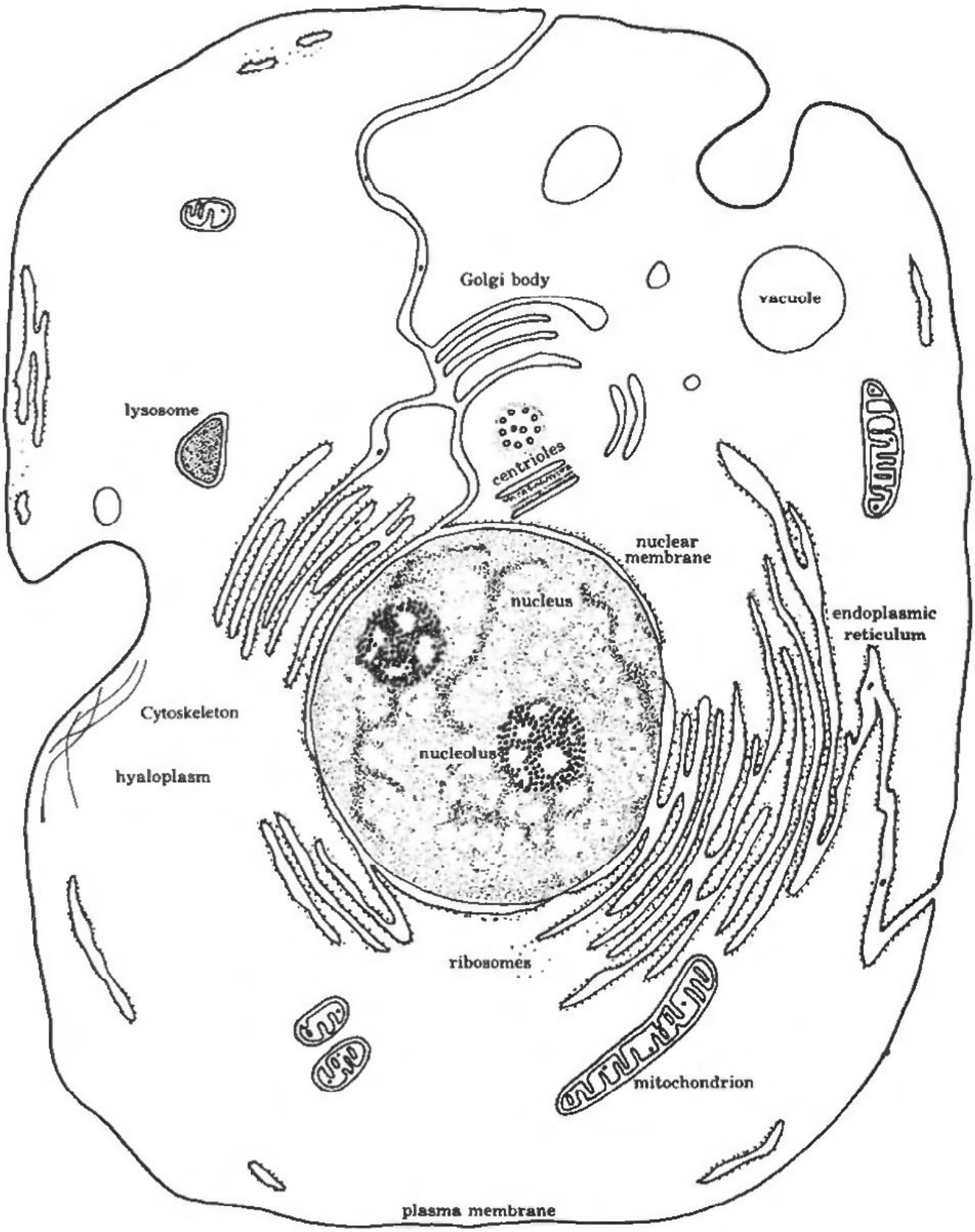


Fig. 1-3 Diagram of a mammalian cell. The organelles are approximately the correct relative sizes.

Question: Can a procedure similar to isopycnic separation in a centrifugal field be used to separate different *macromolecules*?

Yes; in fact one way of preparing and purifying DNA fragments for genetic engineering uses density gradients of CsCl. Various proteins also have different densities and thus can be separated on sucrose density gradients; however, the time required to attain equilibrium is much longer, and higher centrifuge velocities are needed than is the case for organelles.

1.3 SUBCELLULAR ORGANELLES

Question: What does a typical animal cell look like?

There is no such thing as a typical animal cell, since cells vary in overall size, shape, and distribution of the various subcellular organelles. Fig. 1–3 is, however, a composite diagram that indicates the relative sizes of the various *microbodies*.

Plasma Membrane

The *plasma membrane* (Fig. 1-4) is the outer boundary of the cell; it is a continuous sheet of lipid molecules (Chap. 6) arranged as a molecular bilayer 4–5 nm thick. In it are embedded various proteins that function as enzymes (Chap. 8), structural elements, and molecular pumps and selective channels that allow entry of certain small molecules into and out of the cell, as well as receptors for hormones and cell growth factors (Chap. 6).

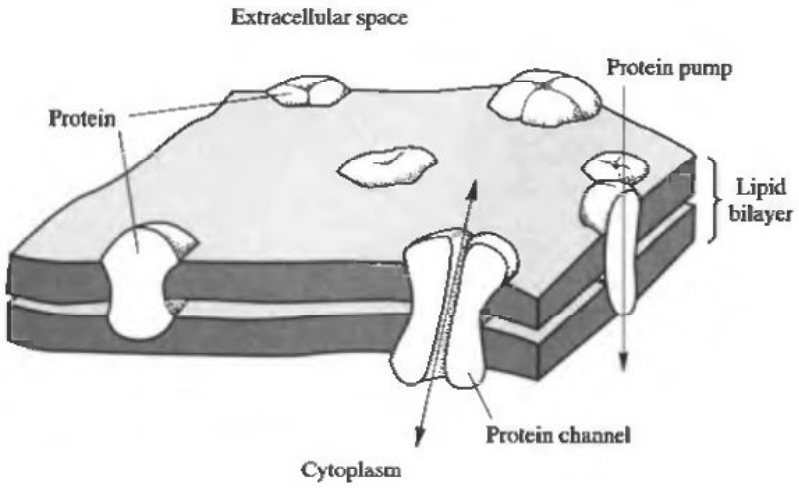
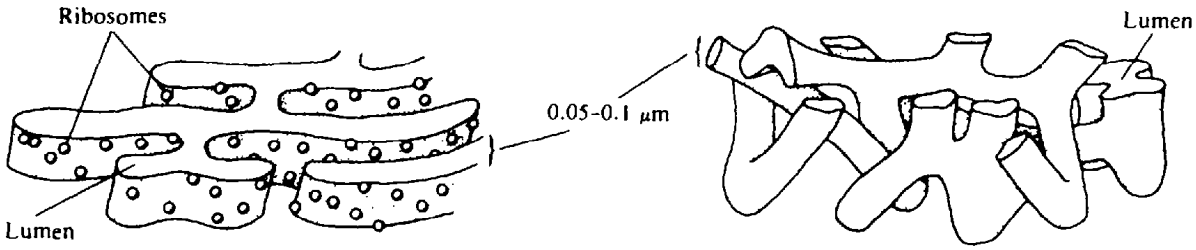


Fig. 1-4 Plasma membrane.

Endoplasmic Reticulum (ER)

The *endoplasmic reticulum* (ER) is composed of flattened sacs and tubes of membranous bilayers that extend throughout the cytoplasm enclosing a large intracellular space. The *luminal* space (Fig. 1-5) is continuous with the outer membrane of the *nuclear envelope* (Fig. 1-10). It is involved in the synthesis of proteins and their transport to the cytoplasmic membrane (via *vesicles*, small spherical particles with an outer bilayer membrane). The *rough* ER (RER) has flattened stacks of membrane that are studded on the outer (cytoplasmic) face with *ribosomes* (discussed later in this section) that



(*a*) Rough endoplasmic reticulum (*b*) Smooth endoplasmic reticulum

Fig. 1-5

actively synthesize proteins (Chap. 17). The *smooth* ER (SER) is more tubular in cross section and lacks ribosomes; it has a major role in lipid metabolism (Chap. 13).

EXAMPLE 1.8

What mass fraction of the lipid membranes of a liver cell is plasma membrane?

Only about 10 percent; the remainder is principally ER and mitochondrial membrane.

Golgi Apparatus

The *Golgi apparatus* is a system of *stacked*, membrane-bound, flattened sacs organized in order of decreasing breadth (see Fig. 1-6). Around this system are small vesicles (50-nm diameter and larger); these are the secretory vacuoles that contain protein that is *released* from the cell (see Example 1.6).

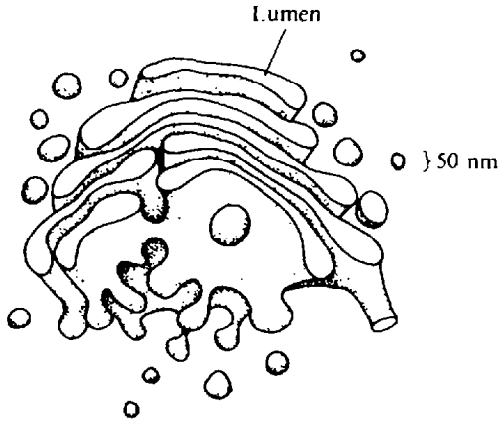


Fig. 1-6 Golgi apparatus and secretory vesicles.

The pathway of secretory proteins and glycoproteins (protein with attached carbohydrate) through *exocrine* (secretory) gland cells in which *secretory vacuoles* are present is well established. However, the exact pathway of exchange of the membranes between the various organelles is less clear and could be either one or a combination of both of the schemes shown in Fig. 1-7.

In the *membrane flow* model of Fig. 1-7, membranes move through the cell from ER → Golgi apparatus → secretory vacuoles → plasma membrane. In the *membrane shuttle* proposal, the vesicles shuttle between ER and Golgi apparatus, while secretory vacuoles shuttle back and forth between the Golgi apparatus and the plasma membrane.

Question: What controls the *directed* flow of membranous organelles?

No one really knows; it is one of the great wonders of cell physiology yet to be fully understood.

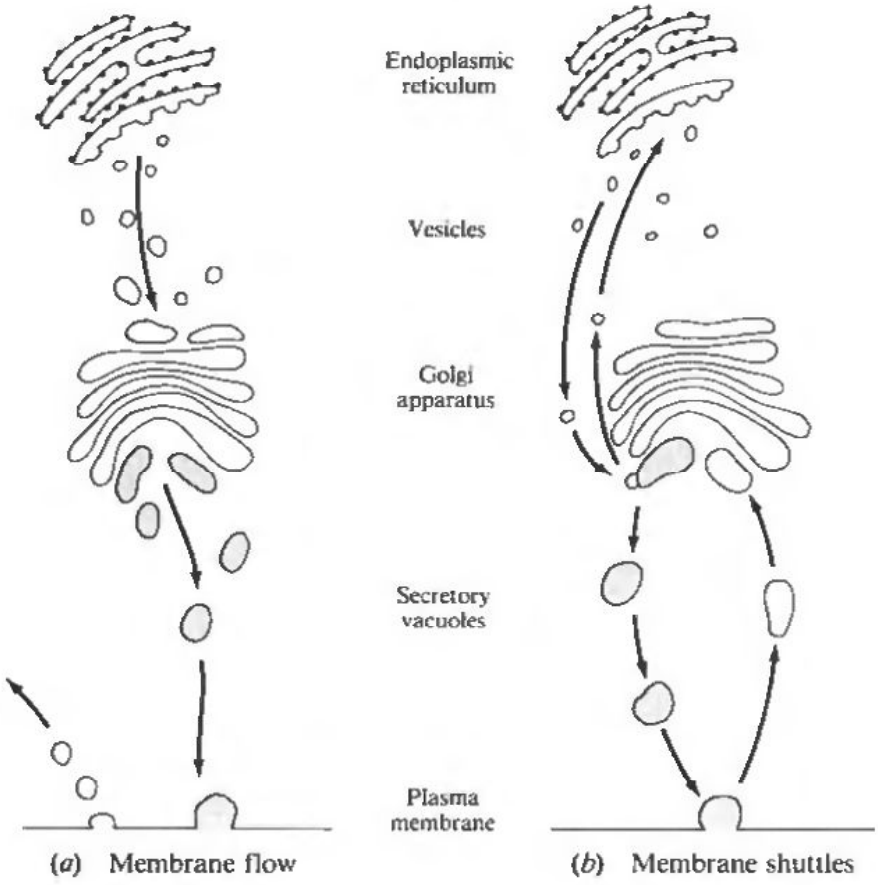


Fig. 1-7 Possible membrane-exchange pathways during secretion of protein from a cell.

Lysosomes

Lysosomes are membrane-bound vesicles that contain *acid hydrolases*; these are enzymes that catalyze hydrolytic reactions and function optimally at a pH (~5) found in these organelles. Lysosomes range in size from 0.2 to 0.5 μm. They are instrumental in intracellular digestion (*autophagy*) and the digestion of material from outside the cell (*heterophagy*). Heterophagy, which is involved with the body's removal of bacteria, begins with the invagination of the plasma membrane, a process called *endocytosis*; the whole digestion pathway is shown in Fig. 1-8.

Since lysosomes are involved in digesting a whole range of biological material, exemplified by the destruction of a whole bacterium with all its different types of macromolecules, it is not surprising to find that a large number of *different* hydrolases reside in lysosomes. These enzymes catalyze the breakdown of nucleic acids, proteins, cell wall carbohydrates, and phospholipid membranes (see Table 1.1).

Mitochondria

Mitochondria are membranous organelles (Fig. 1-9) of great importance in the energy metabolism of the cell; they are the source of most of the ATP (Chap. 14) and the site of many metabolic reactions. Specifically, they contain the enzymes of the citric acid cycle (Chap. 12) and the electron-transport chain (Chap. 14), which includes the main oxygen-utilizing reaction of the cell. A mammalian liver cell contains about 1,000 of these organelles; about 20 percent of the cytoplasmic volume is mitochondrial.

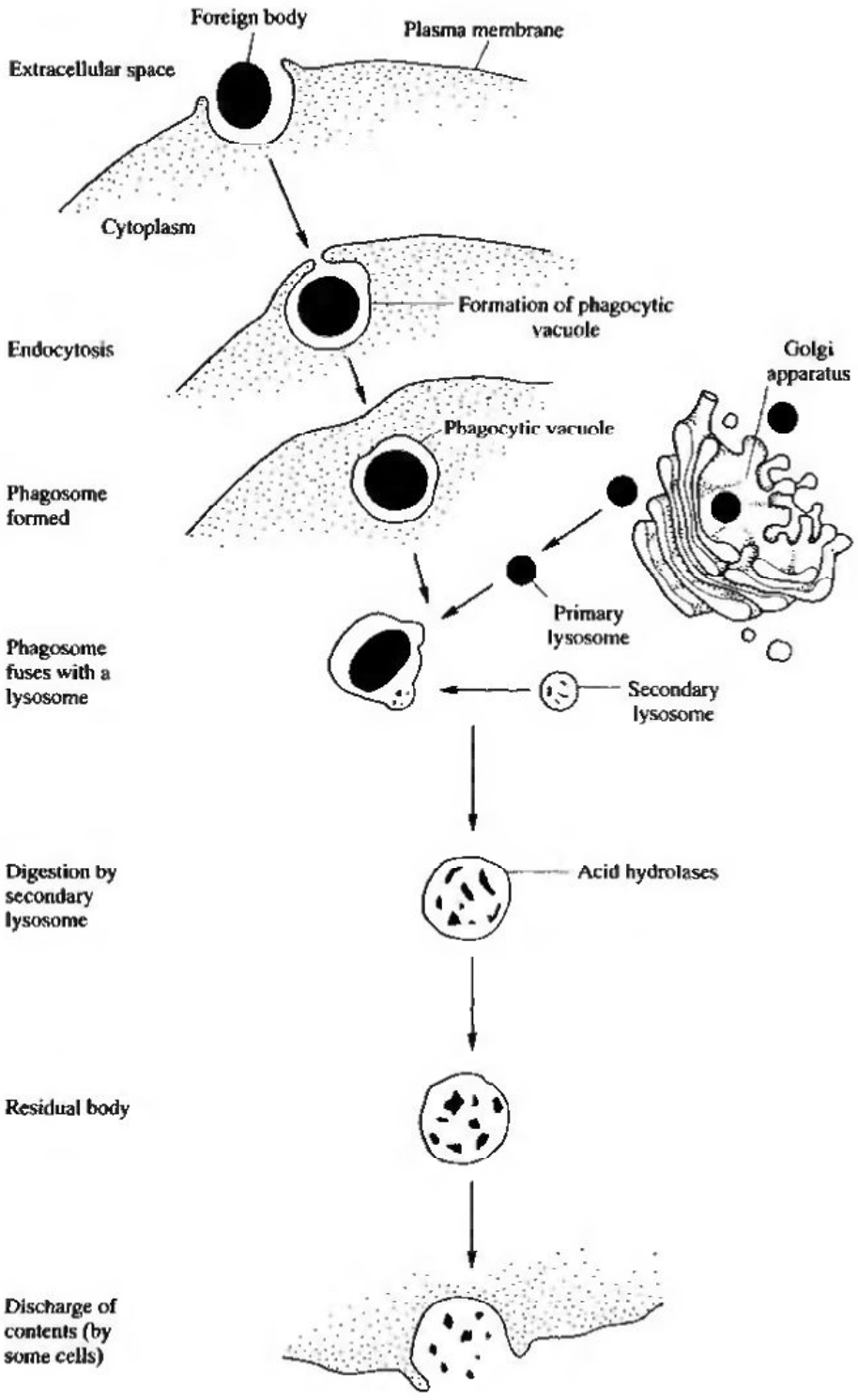


Fig. 1-8 Heterophagy in a mammalian cell, typically a *macrophage*.

Table 1.1. Mammalian Lysosomal Enzymes and Their Substrates

Enzyme	Natural Substrate	Tissue Location
Proteases		
Cathepsin	Most proteins	Most tissues
Collagenase	Collagen (Chap. 4)	Bone
Peptidases	Peptides (Chap. 3)	Most tissues
Lipases		
A range of esterases	Esters of fatty acids (Chap. 13)	Most tissues
Phospholipases	Phospholipids (Chap. 6)	Most tissues
Phosphatases		
Acid phosphatase	Phosphomonoesters (e.g., 2-phosphoglycerol)	Most tissues
Acid phosphodiesterase	Oligonucleotides (Chap. 7)	Most tissues
Nucleases		
Acid ribonuclease	RNA (Chap. 7)	Most tissues
Acid deoxyribonuclease	DNA (Chap. 7)	Most tissues
Polysaccharidases and mucopolysaccharidases		
β-Galactosidase	Galactosides of membranes (Chap. 6)	Liver, brain
α-Glucosidase	Glygogen (Chap. 11)	Macrophages, liver
β-Glucosidase	Glycosphingolipids (Chap. 6)	Brain, liver
β-Glucuronidase	Polysaccharides	Macrophages
Lysozyme	Bacterial cell wall and mucopolysaccharides (Chap. 8)	Kidney
Hyaluronidase	Hyaluronic acid and chrondroitin sulfate (Chap. 2)	Liver
Arylsulfatase	Organic sulfates	Liver, brain

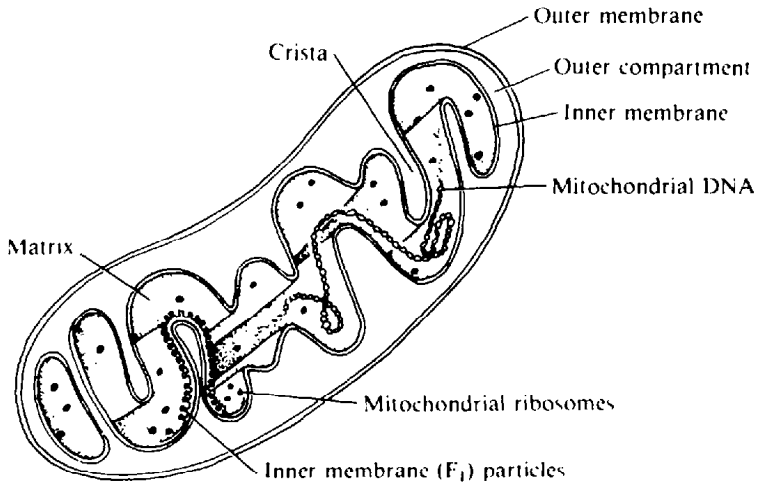


Fig. 1-9 Mitochondrion.

EXAMPLE 1.9

Mitochondria were first observed by R. Altmann in 1890. He named them *bioblasts*, because he speculated that they and chloroplasts (the green chlorophyll-containing organelles of plants) might be intracellular *symbionts* that arose from bacteria and algae, respectively. This idea lay in disrepute until the recent discovery of mitochondrial nucleic acids.

In histology, mitochondria can be stained *supravitally*; i.e., the metabolic activity of the functional (*vital* = living) organelle or cell allows selective staining. The *reduced* form of the dye Janus green B is colorless, but it is *oxidized* by mitochondria to give a light-green pigment that is easily seen in light microscopy.

Mitochondria are about the size of bacteria. They have a diameter of 0.2 to 0.5 μm and are 0.5 to 7 μm long. They are bounded by *two* lipid bilayers, the inner one being highly folded. These folds are called *cristae*. The innermost space of the mitochondrion is called the *matrix*. They have their own DNA in the form of at least one copy of a circular double helix (Chap. 7), about 5 μm in overall diameter; it differs from nuclear DNA in its density and denaturation temperature by virtue of being richer in guanosine and cytosine (Chap. 7). The different density from nuclear DNA allows its separation by isopycnic centrifugation. Mitochondria also have their own type of *ribosomes* that differ from those in the cytoplasm but are similar to those of bacteria.

Most of the enzymes in mitochondria are *imported* from the cytoplasm; the enzyme proteins are largely coded for by *nuclear* DNA (Chap. 17). The enzymes are disposed in various specific regions of the mitochondria (Table 1.2); this has an important bearing on the direction of certain metabolic processes.

Peroxisomes

Peroxisomes are about the same size and shape as lysosomes (0.3 to 1.5 μm in diameter). However they do *not* contain hydrolases; instead, they contain *oxidative* enzymes that generate *hydrogen peroxide* by catalyzing the combination of oxygen with a range of compounds. The various enzymes present in *high* concentration (even to the extent of forming crystals of protein) are (1) urate oxidase (in many animals but not humans); (2) D-amino acid oxidase, Chap. 15; (3) L-amino acid oxidase; and (4) α-hydroxy acid oxidase (includes lactate oxidase). Also, most of the *catalase* in the cell is contained in peroxisomes; this enzyme catalyzes the conversion of hydrogen peroxide, produced in the other reactions, to water and oxygen.

Cytoskeleton

In the cytoplasm, and especially subjacent to the plasma membrane, are networks of protein filaments that stabilize the lipid membrane and thus contribute to the maintenance of cell shape. In cells that grow and divide, such as liver cells, the cytoplasm appears to be organized from a region near the nucleus that contains the cell's pair of *centrioles* (Chap. 5). There are three main types of cytoskeletal filaments: (1) *microtubules*, 25 nm in diameter, composed of organized aggregates of the protein *tubulin* (Chap. 5); (2) *actin* filaments, 7 nm in diameter (Chap. 5); and (3) so-called *intermediate filaments*, 10 nm in diameter (Chap. 5).

Centrioles

Centrioles are a pair of hollow cylinders that are composed of nine triplet tubules of protein (Chap. 5). The members of a pair of centrioles are usually positioned at right angles to each other. Microtubules form the fine weblike protein structure that appears to be attached to the chromosomes during cell division (mitosis); the web is called the *mitotic spindle* and is attached to the ends of the centrioles. While centrioles are thought to function in chromosome segregation during mitosis, it is worth noting that cells of higher plants, which clearly undergo this process, lack centrioles.

Table 1.2. Enzyme Distribution in Mitochondria

Location	Characteristics or Cross-Reference to Discussion
Outer membrane	
Monoamine oxidase	Neurotransmitter; catabolism
Rotenone-insensitive NADH–cytochrome *c* reductase	Chap. 14
Kynurenine hydroxylase	Tryptophan catabolism; Chap. 15
Fatty acid–CoA ligase	Chap. 13
Space between inner and outer membrane	
Adenylate kinase	AMP + ATP 2ADP
Nucleoside diphosphokinase	*X*DP + *Y*TP *X*TP + *Y*DP where *X* and *Y* are any of several ribonucleosides
Inner membrane	
Respiratory chain enzymes	Chap. 14
ATP synthase	Chap. 14
Succinate dehydrogenase	Chap. 14
β-Hydroxybutyrate dehydrogenase	Chap. 13
Carnitine–fatty acid acyltransferase	Chap. 13
Matrix	
Malate and isocitrate dehydrogenase	Chap. 12
Fumarase and aconitase	Chap. 12
Citrate synthase	Chap. 12
2-Oxoacid dehydrogenase	Chap. 12
β-Oxidative enzymes for fatty acids	Chap. 13
Carbamoyl phosphate synthetase I	Chap. 15
Ornithine carbamoyltransferase	Chap. 15

Ribosomes

Ribosomes are the site of protein synthesis and exist: (1) in the cytoplasm as rosette-shaped groups called *polysomes* (in immature red blood cells there are usually five per group); (2) on the outer face of the RER; or (3) in the mitochondrial matrix, although this last type is different in size and shape from ribosomes in the cytoplasm. Ribosomes are composed of RNA and protein and range in size from 15 to 20 nm. Their central role in protein synthesis is described in Chap. 16.

EXAMPLE 1.10

Ribosomes were first isolated by differential centrifugation and then examined by electron microscopy. This and related work by George Palade in the early 1950s earned him the Nobel prize in 1975. For a time ribosomes were known to electron microscopists as *Palade's granules*.

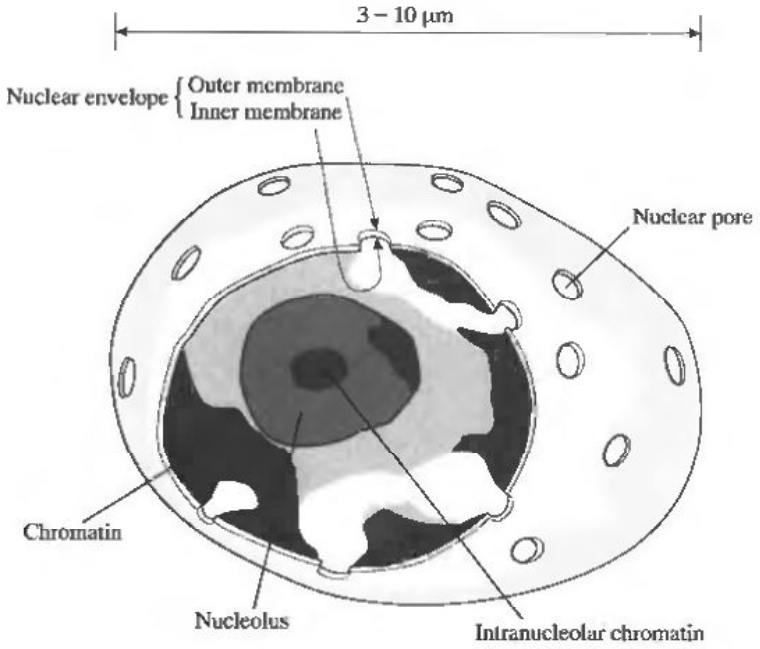


Fig. 1-10 Mammalian cell nucleus.

Nucleus

The *nucleus* is the most conspicuous organelle of the cell (see Fig. 1-10). It is delimited from the cytoplasm by a membranous envelope called the *nuclear membrane*, which actually consists of two membranes forming a flattened sac. The nuclear membrane is perforated by *nuclear pores* (60 nm in diameter), which allow transfer of material between the *nucleoplasm* and the cytoplasm. The nucleus contains the *chromosomes*, which consist of DNA packaged into *chromatin* fibers by association of the DNA with an equal mass of *histone* proteins (Chap. 16).

Nucleolus

The *nucleolus* is composed of 5 to 10 percent RNA, and the remainder of the mass is protein and DNA. In light microscopy it appears to be spherical and *basophilic* (Prob. 1.1). Its function is the synthesis of ribosomal RNA (Chap. 17). There may be more than one per nucleus.

Chromosomes

Chromosomes are the bearers of the hereditary instructions in a cell; thus they are the overall regulators of cellular processes. Important features to note about chromosomes are:

(*a*) *Chromosome number*. In animals, each *somatic* cell (body cells, excluding sex cells) contains one set of chromosomes inherited from the female parent and a comparable (*homologous*) set from the male parent. The number of chromosomes in the dual set is called the *diploid number*; the suffix *-ploid* means "a set" and the *di* refers to the multiplicity of the set (in this case, "two"). Sex cells (called *gametes*) contain *half* as many chromosomes as found in somatic cells and are therefore referred to as *haploid* cells. A *genome* is the set of chromosomes that corresponds to the haploid set of a species.

EXAMPLE 1.11

Human somatic cells contain 46 chromosomes, cattle 60, and fruit fly 8. Thus, the diploid number bears no relationship to the species' positions in the phylogenetic scheme of classification.

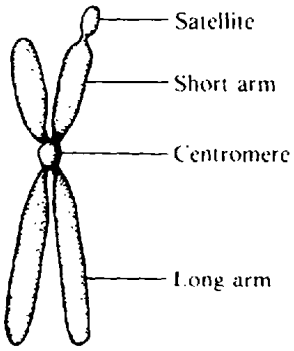


Fig. 1-11 Mammalian chromosome.

(*b*) *Chromosome morphology.* Chromosomes become visible under the light microscope only at certain phases of the nuclear division cycle. Each chromosome in the genome can usually be distinguished from the others by such features as: (1) relative length of the whole chromosome; (2) the position of the *centromere*, a structure that divides the chromosome into a crosslike structure with two pairs of arms of different length; (3) the presence of knobs of chromatin called *chromomeres*; and (4) the presence of small terminal extensions called *satellites* (Fig. 1-11).

EXAMPLE 1.12

In the clinical investigation of infants or fetuses with possible inborn errors of metabolism or morphology, it is common practice to prepare a *karyotype*. Usually, white cells are cultured and then stimulated to divide. The predivision cells are squashed between glass slides, causing the cellular nuclei to disgorge their chromosomes, which are then stained with a blue dye. The chromosomes are photographed and then ordered according to their length, the longest pair being numbered 1. The sex chromosomes do not have a number.

The inherited disorder *Down syndrome* (also called mongolism) involves mental retardation and distinctive facial features. It results from the inclusion of an extra chromosome 21 in each somatic cell of the body. Hence, the condition is called *trisomy 21*.

(*c*) *Autosomes and sex chromosomes.* In humans, gender is associated with a morphologically dissimilar pair of chromosomes called the *sex chromosomes*. The two members of the pair are labeled *X* and *Y*, *X* being the larger. Genetic factors on the *Y* chromosome determine maleness. All chromosomes, exclusive of the sex chromosomes, are called *autosomes*.

1.4 CELL TYPES

There are over 200 different cell types in the human body. These are arranged in a variety of different ways, often with mixtures of cell types, to form *tissues*. Among this vast array of types are some highly specialized ones.

Red Blood Cell (Erythrocyte)

Erythrocytes are small compared with most other cells and are peculiar because of their biconcave disk shape (see Fig. 1-12). They have no nucleus, because it is extruded just before the release of the cell into the blood stream from the bone marrow, where the cells develop. Their cytoplasm has no organelles and is full of the protein *hemoglobin* that binds O_2 and CO_2. In the cytoplasm are other proteins also, namely, (1) the submembrane cytoskeleton, (2) enzymes of the glycolytic and

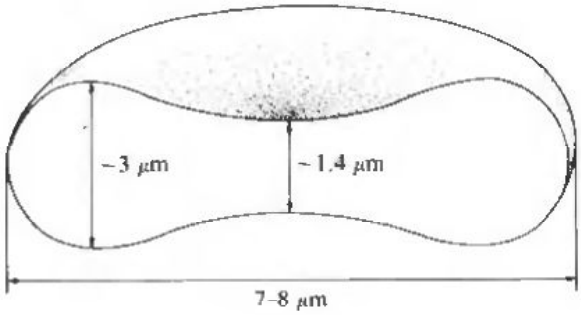


Fig. 1-12 Human erythrocyte.

pentose phosphate pathways (Chap. 11), and (3) a range of other hydrolytic and special-function enzymes that will not be discussed here. In the membrane are specialized proteins associated with (1) *anion transport*, and (2) the bearing of the carbohydrate cell-surface antigens (blood group substances).

Adipocyte

Adipocytes are the specialized cells of fat tissue (Fig. 1-13). The cells range in size from 60 to 120 μm in diameter and have the characteristic feature of a huge vacuole that is full of triglycerides (Chap. 6). The nucleus and mitochondria are flattened on one inner surface of the plasma membrane, and there is only a small amount of endoplasmic reticulum.

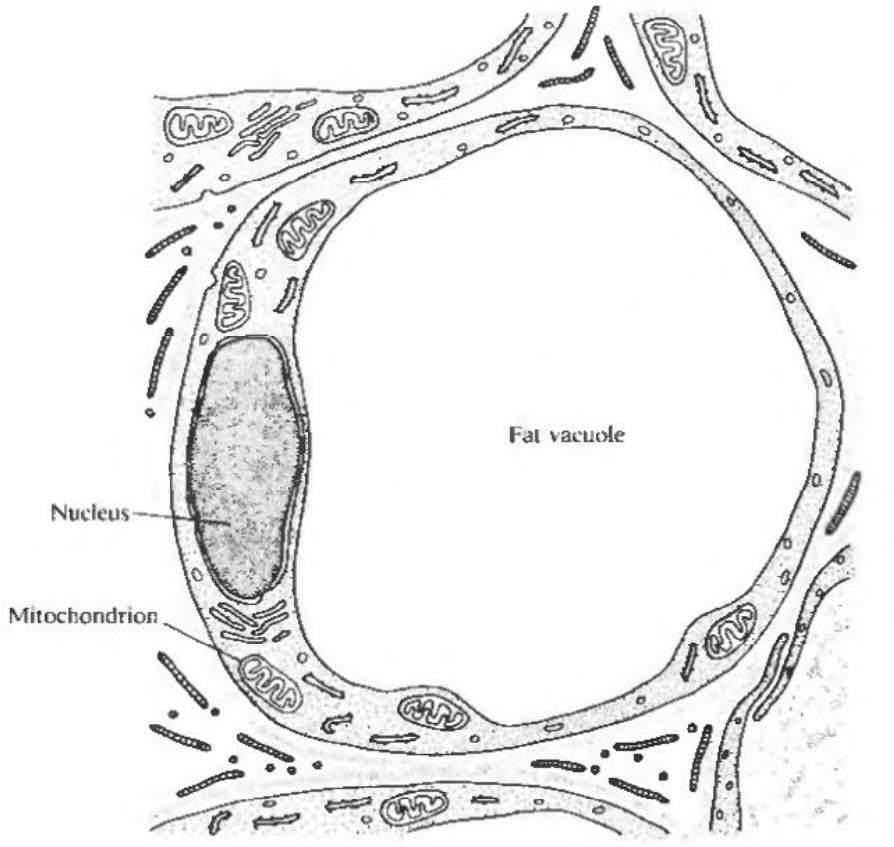


Fig. 1-13 Adipocyte.

Liver Cell (Hepatocyte)

Liver tissue contains an array of cell types, but the preponderant one is the *hepatocyte*. It has an overall structure much like that of the cell in Fig. 1-3. The cells are arranged in long, branching columns of about 20 cells in a cross section around a central *bile cannaliculus* (channel). Into the cannaliculus the cells secrete bile. The liver is the main producer of urea (Chap. 15), stores glycogen (Chap. 11), synthesizes many of the amino acids used by other tissues (Chap. 3), and produces serum proteins, among many other metabolic roles.

Muscle Cell (Myocyte)

Muscle cells produce mechanical force by contraction. In vertebrates there are three basic types:

1. *Skeletal muscle* moves the bones attached to joints. These muscles are composed of bundles of long, multinucleated cells. The cytoplasm contains a high concentration of a special macromolecular contractile-protein complex, *actomyosin* (Chap. 5). There is also an elaborate membranous network called the *sarcoplasmic reticulum* that has a high Ca^{2+} content. The contractile-protein complex has a banded appearance under microscopy.
2. *Smooth muscle* is the type in the walls of blood vessels and the intestine. The cells are long and spindle-shaped, and they lack the banding of skeletal muscle cells.
3. *Cardiac muscle* is the main tissue of the heart. The cells are similar in appearance to those of skeletal muscle but in fact have a different biochemical makeup.

Epithelia

Epithelial cells (Fig. 1-14) form the coherent sheets that line the inner and outer surfaces of the body. There are many specialized types, but the main groups are as follows:

1. *Absorptive cells* have numerous hairlike projections called *microvilli* on their *outer* surface; they increase the surface area for absorption of nutrients from the gut lumen and other areas.
2. *Ciliated cells* have small membranous projections (*cilia*) that contain interior contractile proteins; cilia beat in synchrony and serve to sweep away foreign particles on the surface of the respiratory tract, i.e., in the lungs and the nasal lining.
3. *Secretory cells* occur in most epithelial surfaces; e.g., sweat gland cells in the skin and mucus-secreting cells in the intestine and respiratory tract.

1.5 THE STRUCTURAL HIERARCHY IN CELLS

The organic molecules that are building blocks of biological macromolecules are very small; e.g., the amino acid alanine is only 0.7 nm long, whereas a typical globular protein, hemoglobin (Chap. 5), which consists of 574 amino acids, has a diameter of ~6 nm. In turn, protein molecules are small compared with the ribosomes that synthesize them (Chap. 17); these macromolecular aggregates are composed of over 70 different proteins and four nucleic acid strands. They have a molecular weight (M_r) of around 2.8×10^6 and a diameter of ~20 nm. In contrast, mitochondria contain their own ribosomes and DNA and range in length up to 7 μm. Intracellular vesicles are often seen to be about the same size as mitochondria, and yet the Golgi apparatus or the lipid vacuole of an adipocyte is much larger. The nucleus may be even larger and also contains some ribosomes and other macromolecular aggregates, including, most importantly, the chromosomes.

Even though the building blocks of macromolecules are small in relation to the size of the cell (e.g., the ratio of the volume of one molecule of alanine to that of the red blood cell is ~$1:10^{11}$),

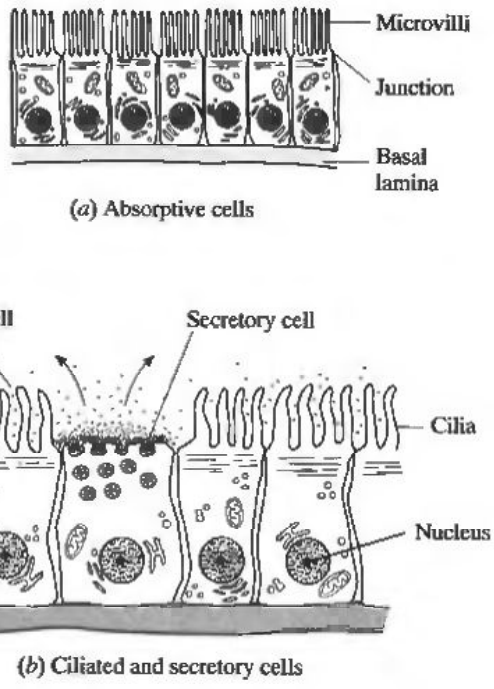


Fig. 1-14 Epithelial cells.

a defect in one amino acid in the sequence of a protein can profoundly affect not only the protein but also the cell structure. Furthermore, an altered enzymatic activity or binding affinity can greatly influence the survival of not only the cell but the whole being.

EXAMPLE 1.13

In humans having the inherited disease called *sickle-cell anemia*, the hemoglobin molecules of the erythrocytes are defective; 2 of the 574 amino acids in the protein are substituted for another. Specifically, glutamate in position 6 of each of the two β chains of the hemoglobin tetramer (see Chap. 5) is replaced by valine. This single change increases the likelihood of the molecules aggregating when they are deoxygenated. The aggregated protein forms large *paracrystalline structures* (called *tactoids*) inside the cell and distorts it into a relatively inflexible sickle shape. These cells tend to clog small blood vessels and capillaries and lead to poor oxygen supply in many organs. Also, they are more fragile and thus rupture, reducing the number of cells and causing anemia.

Solved Problems

METHODS OF STUDYING THE STRUCTURE AND FUNCTION OF CELLS

1.1. Basic dyes such as methylene blue or toluidine blue are positively charged at the pH of most staining solutions used in histology. Thus the dyes bind to acidic (i.e., those that become negatively charged on dissociation of a proton) substances in the cell. These acidic molecules are therefore referred to as basophilic substances. Give some examples of basophilic substances.

SOLUTION

Examples of basophilic cell components are DNA and RNA; the latter includes messenger RNA (Chap. 7) and ribosomes. The youngest red blood cells in the blood circulation contain a basophilic *reticulum* (network) in their cytoplasm; this is composed of messenger and ribosomal RNA. The network slowly dissolves over the first 24 hours of the cell's life in the circulation. This readily identifiable red blood cell type is called the *reticulocyte*.

1.2. Acidic dyes such as eosin and acid fuchsin have a net negative charge at the pH of usual staining solutions. Therefore, they bind to many cellular proteins that have a net positive charge. Give some regions of a liver cell that might be *acidophilic*.

SOLUTION

The cytoplasm, mitochondrial matrix, and the inside of the smooth endoplasmic reticulum are acidophilic; all these regions almost exclusively contain protein.

1.3. Describe a possible means for the cytochemical detection and localization of the enzyme *glucose 6-phosphatase*; it exists in liver and catalyzes the following reaction:

$$\text{Glucose 6-phosphate} \xrightarrow{H_2O} \text{Glucose} + \text{Phosphate}$$

SOLUTION

Incubate a tissue slice at 37°C with glucose 6-phosphate in a suitable buffer solution. Wash the tissue free of the substrate, and then precipitate the phosphate ions by the addition of lead nitrate to the tissue slice. The remainder of the preparation is as described in Example 1.5. In liver cells the reaction product is found *within* the endoplasmic reticulum, thus indicating the location of the enzyme.

1.4. How may cells be disrupted in order to obtain subcellular organelles by centrifugal fractionation?

SOLUTION

There are several ways of disrupting cells:

1. *Osmotic lysis*: The plasma membranes of cells are water-permeable but are impermeable to large molecules and some ions. Thus, if cells are placed into water or dilute buffer, they swell owing to the *osmotically* driven influx of water. Since the plasma membrane is not able to stretch very much (the red blood cell membrane can stretch only up to 15 percent of its normal area before disruption), the cells burst. The method is effective for isolated cells but is not so effective for tissues.
2. *Homogenizers*: One of these is described in Example 1.7.
3. *Sonication*: This involves the generation of shear forces in a cell sample in the vicinity of a titanium probe (0.5 mm in diameter and 10 cm long) that vibrates at ~20,000 Hz. The device contains a crystal of lead zirconate titanate that is *piezoelectric*, i.e., it expands and contracts when an oscillatory electric field is applied to it from an electronic oscillator. The ultrasonic pressure waves cause *microcavitation* in the sample, and this disrupts the cell membranes, usually in a few seconds.

SUBCELLULAR ORGANELLES

1.5. On the basis of the pathway of heterophagy (Fig. 1-8), make a proposal for the pathway of *autophagic* degradation of a mitochondrion.

SOLUTION

Figure 1-15 shows the scheme for autophagic degradation of a mitochondrion. Note that once the so-called *phagosome* has been formed, the process of digestion, etc., is the same as for heterophagy (Fig. 1-8).

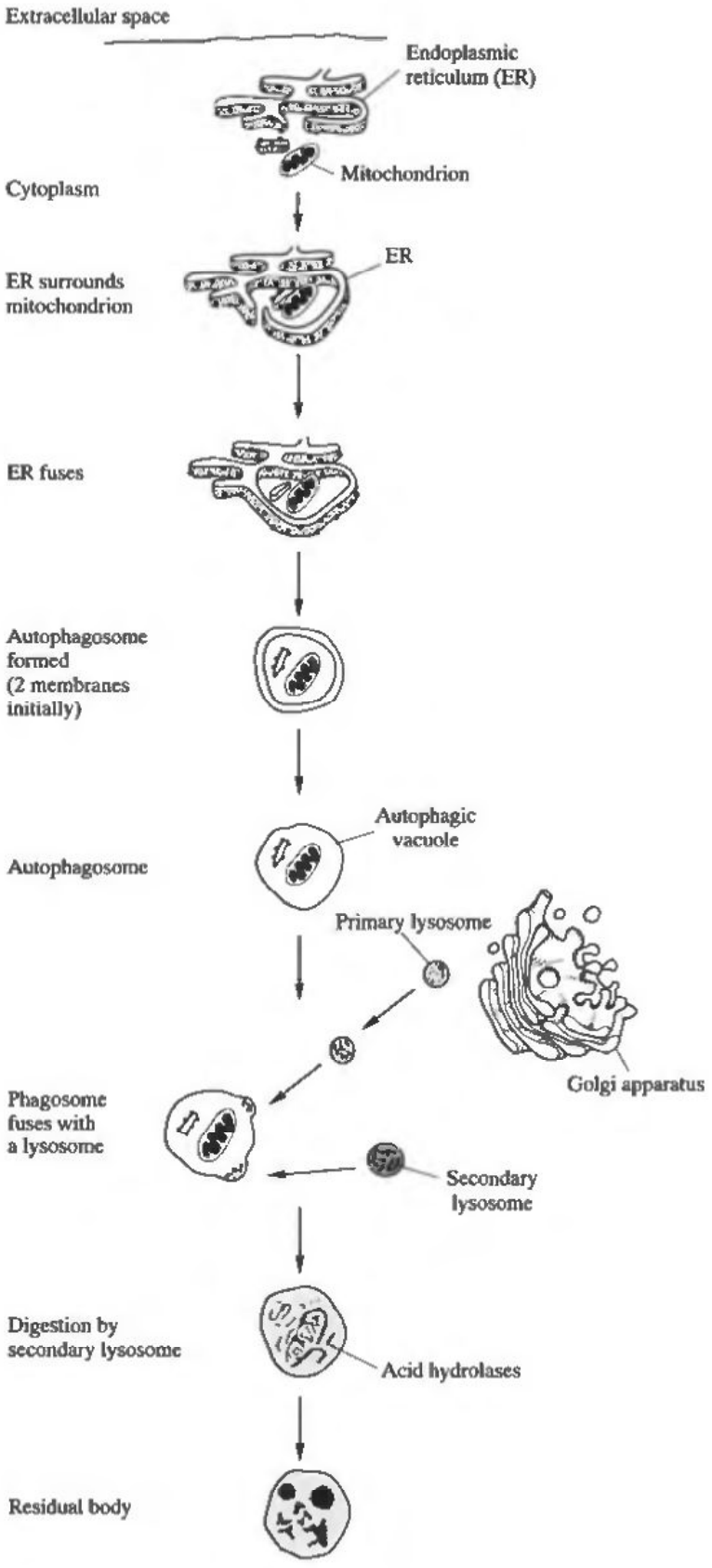


Fig. 1-15 The process of autophagy of a mitochondrion.

1.6. There is an inherited disease in which a person's lysosomes lack the enzyme *β-glucosidase* (Table 1.1). What are the clinical and biochemical consequences of this deficiency?

SOLUTION

The disease is called *Gaucher disease* and is the most common of the so-called *sphingolipidoses*; its incidence in the general population is ~1:2,500. This class of disease results from defective hydrolysis of membrane components called *glycosphingolipids* (Chap. 6) that are normally turned over in the cell by hydrolytic breakdown in the lysosomes. The glycosphingolipids are lipid molecules with attached carbohydrate groups. An inability to remove glucose from these molecules results in their accumulation in the lysosomes. In fact over a few years, the cells that have rapid membrane turnover, such as in the liver and spleen, become engorged with this lipid-breakdown product. Clinically, patients have an enlarged liver and spleen and may show signs of mental deterioration if much of the lipid accumulates in the brain as well.

CELL TYPES

1.7. How many red blood cells are there in an average 70-kg person?

SOLUTION

There are $\sim 3.3 \times 10^{13}$ red blood cells in an average person. The total blood volume is ~6 L, and half of that is red blood cells; i.e., there is ~3 L of red blood cells. Since each cell has a volume of $\sim 90 \times 10^{-15}$ L (Fig. 1-12), the result follows from dividing 3 L by this number.

1.8. If the average life span of a human red blood cell is 120 days, how many red blood cells are produced in an average 70-kg person every second?

SOLUTION

There are 3.2 million red blood cells produced every second! The number produced per second is simply given by the answer from Prob. 1.7, above, divided by 120 days expressed in seconds.

1.9. A *macrophage* is a cell type that is involved in the engulfing of foreign material, such as bacteria and damaged host cells. In view of this specialized phagocytic function, draw what you think an electron microscopist would see in a cross section of the cell.

SOLUTION

The key features of a macrophage are its large system of *lysosomes* and *invaginations of the cytoplasmic membrane* (Fig. 1-16). Also, there is a rich rough endoplasmic reticulum where the lysosomal hydrolytic enzymes are produced. Mitochondria are abundant since the highly active protein synthesis is very demanding of ATP (Chap. 17).

1.10. PAS staining of microscope sections of red blood cells gives a pink stain on one side only of the cell membrane. Which side is it, the extracellular or the intracellular side?

SOLUTION

The extracellular side is stained pink. All glycoproteins and glycolipids of the plasma membrane of red blood cells and *all* other cells are on the outside of the cell. No oligosaccharides are present on the inner face of the plasma membrane.

1.11. Why do the vesicles of some *mast* cells, which contain large quantities of *histamine*, stain red with the dye eosin?

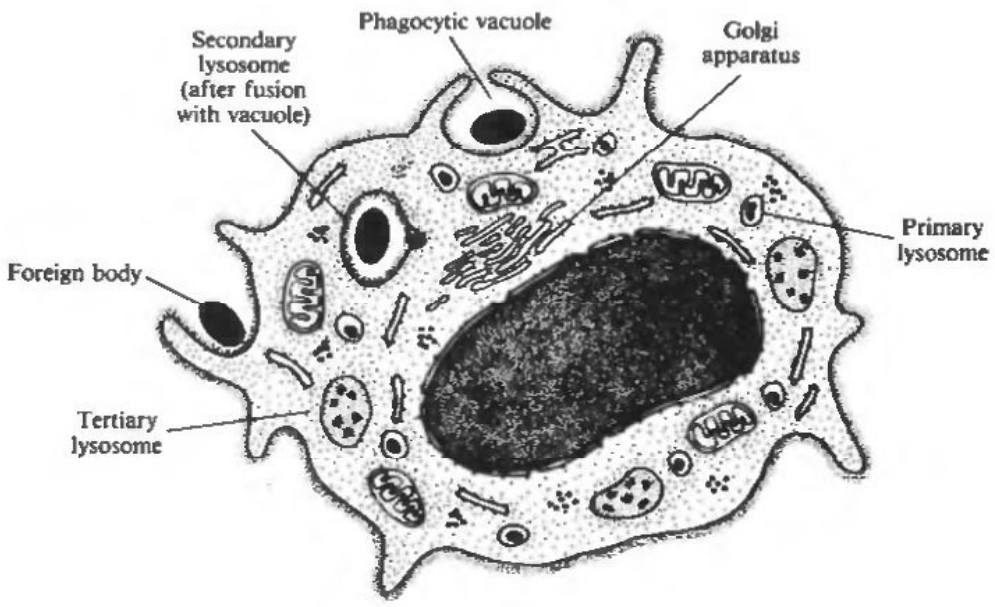
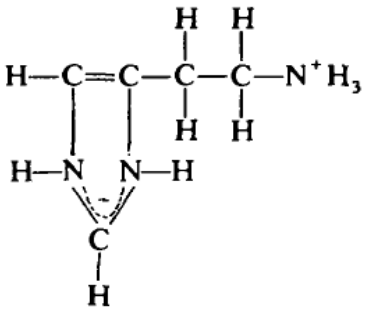


Fig. 1-16 Macrophage.

SOLUTION

Eosin is negatively charged, and histamine has the structure



Histamine

i.e., it has a maximum charge of +2. The two types of molecules interact electrostatically inside the vesicles, thus the eosin stains the vesicles red.

THE STRUCTURAL HIERARCHY IN CELLS

1.12. The concentration of hemoglobin in human red blood cells is normally $330\,g\,L^{-1}$. The molecular weight (M_r) of hemoglobin is 64,500, and the volume of a red cell is ~90 fL. How many molecules of hemoglobin are there in one human red blood cell?

SOLUTION

There are $\sim 3 \times 10^8$ molecules of hemoglobin in one erythrocyte. The number of moles of hemoglobin in one cell is

$$\frac{330 \times 90 \times 10^{-15}}{64{,}500} = 4.6 \times 10^{-16}$$

Since Avogadro's number is the number of molecules per mole of a compound, the previous number is multiplied by Avogadro's number to give the required estimate:

$$4.6 \times 10^{-16} \times 6.02 \times 10^{23} \simeq 3 \times 10^8$$

1.13. The mean generation time of a red cell, from the stem cell to a mature reticulocyte, is ~90 h. The phase in the cell generation pathway in which most of the hemoglobin is synthesized is ~40 h. How many hemoglobin molecules are synthesized per human red blood cell per second?

SOLUTION

Since, from Prob. 1.12, we saw that the cell contains $\sim 3 \times 10^8$ hemoglobin molecules, we proceed by simply dividing this number by the time taken to generate them, 40 h. This gives the rate of production, namely, ~2,000 molecules per second.

1.14. It has been estimated that it takes ~1 min to synthesize one hemoglobin subunit from its constituent amino acids. Using this fact, calculate how many hemoglobin molecules are produced on average at any one time in the differentiation of the cell.

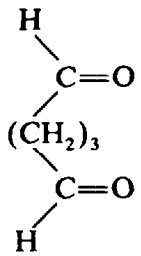
SOLUTION

From Prob. 1.13, ~2,000 hemoglobin molecules are produced per second; this is equal to $\sim 1.2 \times 10^5$ per minute. However, hemoglobin is a *tetrameric* protein (four subunits; Chap. 5), so four times 1.2×10^5 chains are produced per minute: 4.8×10^5.

Supplementary Problems

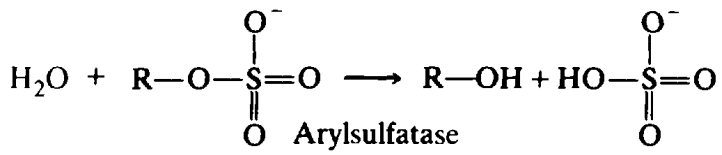
1.15. A commonly used test of the viability of cells in tissue culture is whether or not they exclude a so-called supravital dye such as toluidine blue. If the cells exclude the dye, they are considered to be viable. What is the biochemical basis of this test?

1.16. The chemical compound *glutaraldehyde* has the structure



It is used as a fixative of tissues for light and electron microscopy. What chemical reaction is involved in this fixation process?

1.17. Outline the design of a histochemical procedure for the localization of the enzyme arylsulfatase in tissues; the enzyme catalyzes the following reaction type:



1.18. In an attempt to determine the localization of glycogen in the liver, could there be any problems of interpretation of the electron microscopic autoradiographic images if [^{3}H]glucose were used as the radioactive precursor molecule of glycogen?

1.19. *Microsomes* are small, spherical, membranous vesicles with attached ribosomes. During differential sedimentation, they sediment only in the late stages of a preparation, when very high centrifugal forces are used. They don't appear in electron micrographs of a cell. From whence do they arise?

1.20. There are two forms of the enzyme carbamoyl phosphate synthetase, one in the mitochondrial matrix and the other in the cytoplasm. What might be the consequence and role of this so-called *compartmentation* of enzymes?

1.21. Human reticulocytes (Prob. 1.1) continue to synthesize hemoglobin for approximately 24 hours after release into the circulation. Design an electron microscopic experiment using autoradiography to identify *which* of the cells are actively synthesizing the protein.

1.22. (*a*) Who is the *primary* source of the DNA in your mitochondria—your mother or your father?
(*b*) Speculate on possible inheritance patterns if there were a defect in one or the other parent's mitochondria.

1.23. Given that mitochondria do not have the same aggressive autolytic capacity as lysosomes, what might be the significance of having such a complex membranous structure? After all, the endoplasmic reticulum and the plasma membrane could potentially support those enzymes found in mitochondrial membranes.

1.24. The disease *epidermolysis bullosa* involves severe skin ulceration and even loss of the ends of the ears, nose, and fingers. It is possibly the result of a primary defect in the stability of lysosomal membranes.

(*a*) How does this lead to the signs, just mentioned, of the disease?

(*b*) What biochemical procedure might you suggest to treat the disorder?

1.25. In some sufferers of Down syndrome, the somatic cell nuclei do not contain three chromosomes 21. However, there is a chromosomal defect relating to chromosome 21; what might it be?

Chapter 2

Carbohydrates

2.1 INTRODUCTION AND DEFINITIONS

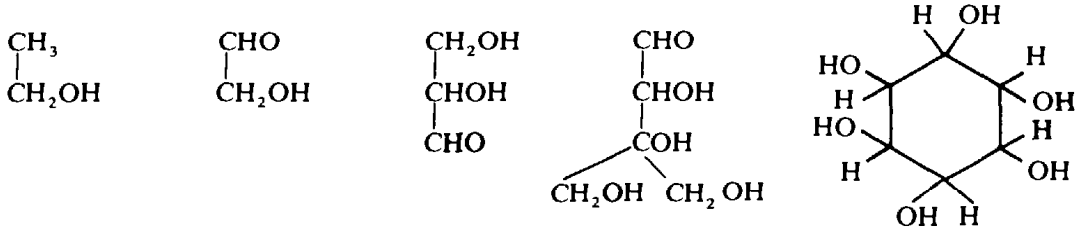
It is not possible to give a simple definition of the term *carbohydrate*. The name was applied originally to a group of compounds containing C, H, and O that gave an analysis of $(CH_2O)_n$, i.e., compounds in which n carbon atoms appeared to be hydrated with n water molecules. These compounds possessed reducing properties because they contained a carbonyl group as either an aldehyde or a ketone, as well as an abundance of hydroxyl groups. The use of the term *carbohydrate* was extended to describe the derivatives of these simple compounds, although the derivatives failed to give the simple analysis shown above. Moreover, many naturally occurring compounds proved to be derivatives in which the reducing group (aldehyde or ketone) had undergone reaction.

The simplest definition of carbohydrates that can be given is that they are *polyhydroxy-aldehydes* or *-ketones*, or compounds derived from these. They range in M_r from less than 100 to well over 10^6. The smaller compounds, containing three to nine carbon atoms, are called *monosaccharides*. The larger compounds are formed by condensation of the smaller ones via *glycosidic bonds*. A *disaccharide* consists of two monosaccharides linked by a single glycosidic bond; a *trisaccharide* is three monosaccharides linked by two glycosidic bonds, etc. *Oligo-* and *polysaccharides* are terms describing carbohydrates with few and many monosaccharide units, respectively.

Because many mono- and oligosaccharides have a sweet taste, carbohydrates of low M_r are often called *sugars*.

EXAMPLE 2.1

Of the compounds shown, (1) and (2) are not carbohydrates because they have only one hydroxyl group each; (3), (4), and (5) are carbohydrates because they have the general formula $(CH_2O)_n$, and are polyhydroxylic.

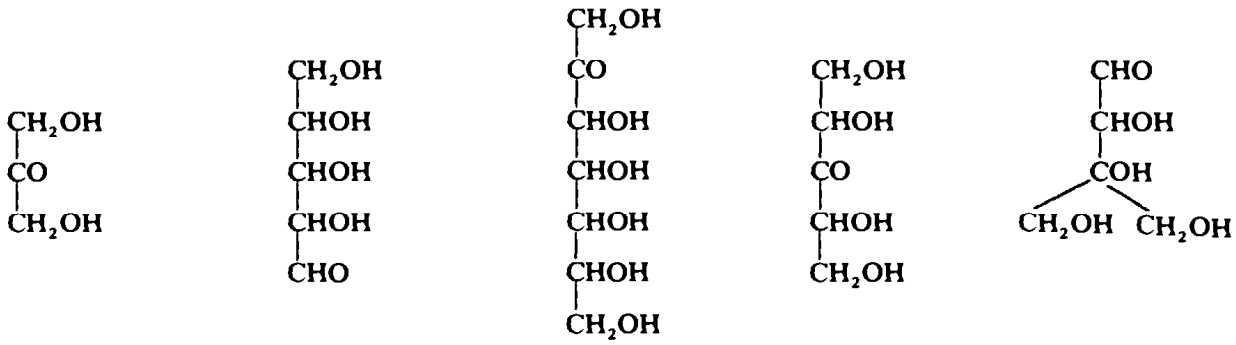


(1) (2) (3) (4) (5)

Let us for now restrict discussion to simple monosaccharides, that is, polyhydroxy compounds containing a carbonyl functional group. There are two series—*aldoses*, containing an aldehyde group, and *ketoses*, containing a ketone group. Simple monosaccharides can also be classified according to the number of carbon atoms they contain—*trioses*, *tetroses*, *pentoses*, *hexoses*, etc., containing three, four, five, and six carbon atoms, respectively. The two systems can be combined. Thus, glucose, the most common sugar, is an *aldohexose*; i.e., a six-carbon monosaccharide with an aldehyde group.

EXAMPLE 2.2

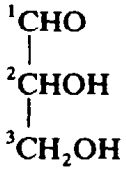
The classification of each of the following monosaccharides is given below the structure.



Ketotriose Aldopentose Ketoheptose Ketopentose Aldopentose

2.2 GLYCERALDEHYDE

The simplest aldose is glyceraldehyde:

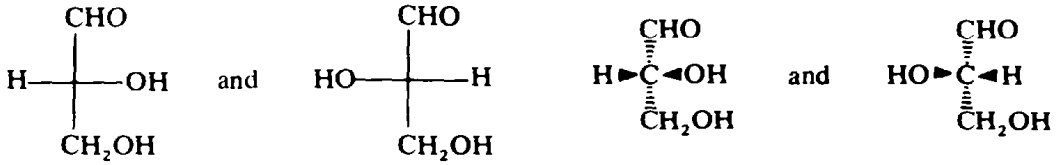


It has reducing properties because it is an aldehyde. The C-2 of glyceraldehyde is a *chiral center* (also known as an asymmetric center), i.e., there are two possible isomers, known as *enantiomers*, of glyceraldehyde.

EXAMPLE 2.3

The structures of the two enantiomers of glyceraldehyde are

These structures, when written as shown below (left), are called *Fischer projection formulas*, which are attempts to represent three-dimensional molecules in two dimensions. For example, the two molecules would appear in three-dimensional space as shown below (right)



where the C-2 is in the plane of the paper, the aldehyde and hydroxymethyl groups are behind the paper (as denoted by the dashed bond lines), and the hydrogen and hydroxyl groups are in front of the paper (as denoted by the solid bond lines). Enantiomers are *mirror images* of each other: when placed in front of a plane mirror, one structure will give an image that is identical with the structure of the other.

In the pairs of figures above, the structure on the left is called D-glyceraldehyde, and that on the right L-glyceraldehyde. The prefixes D- and L- refer to the overall shape of the molecules; more specifically the letters refer to the *configuration*, or arrangement, of groups around the chiral center.

Generally, with compounds containing a single chiral carbon atom, there are only minor differences in the physical and chemical properties of the pure enantiomers. There is one physical property, however, in which enantiomers are markedly different—the property of *optical activity*. This refers to the ability of a solution of an enantiomer to rotate the plane of plane-polarized light. One of a pair of enantiomers will rotate the plane in a *clockwise* direction and is given the symbol (+). The other will rotate the plane in a *counterclockwise* direction and is given the symbol (−). The D enantiomer of glyceraldehyde is (+) and is described more fully as D-(+)-glyceraldehyde; the other is L-(−)-glyceraldehyde. Mixtures of D and L enantiomers will have a net rotation depending on the proportions of the enantiomers; equal proportions give a net rotation of zero, in which case the solution is said to be *racemic*.

Optical activity is measured in a *polarimeter*. The magnitude of the optical activity is measured as an angle of rotation, given the symbol α. The units are degrees or radians (SI).

Question: On what factors will the α of a solution of an optically active compound depend?

The factors are the concentration of the compound, the length of the cell in which the solution is placed, the wavelength of the polarized light, the temperature, and the solvent.

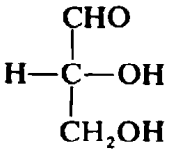
Because of the dependencies noted above, the experimentally measured α is always converted to and expressed as the *specific rotation* $[\alpha]_D^T$, where the super- and subscripts refer, respectively, to the temperature and the wavelength of the light (D referring to the D lines of sodium vapor, 589.2 nm).

$$[\alpha]_D^T = \frac{\alpha}{\text{length of cell (dm)} \times \text{concentration (g cm}^{-3})} \tag{2.1}$$

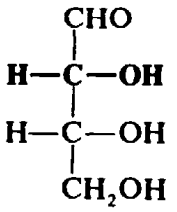
The name of the solvent is given in parentheses after the value is given; e.g., $[\alpha]_D^{25} = +17.5°$ (in water).

2.3 SIMPLE ALDOSES

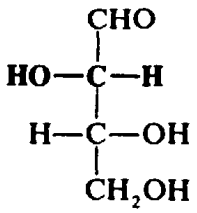
The simple aldoses are related to D- and L-glyceraldehyde in that structurally they may be considered to be derived from glyceraldehyde by the introduction of hydroxylated chiral carbon atoms between C-1 and C-2 of the glyceraldehyde molecule. Thus, two tetroses result when CHOH is introduced into D-glyceraldehyde:



D-Glyceraldehyde



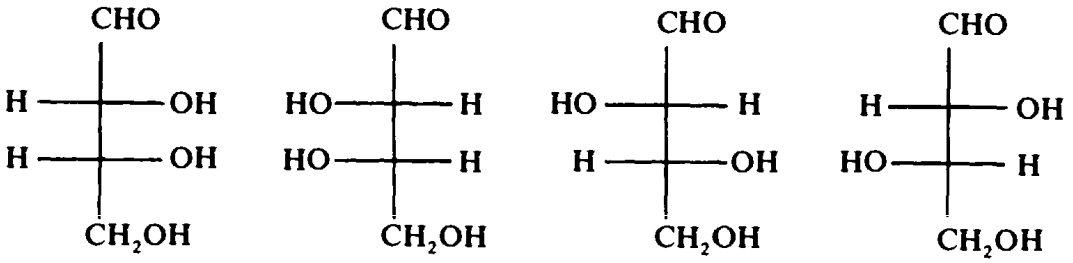
D-Erythrose



D-Threose

Question: Two tetroses can be formed from L-glyceraldehyde. These tetroses are called L-*erythrose* and L-*threose*. Why is it unnecessary to invent new names for the tetroses derived from L-glyceraldehyde?

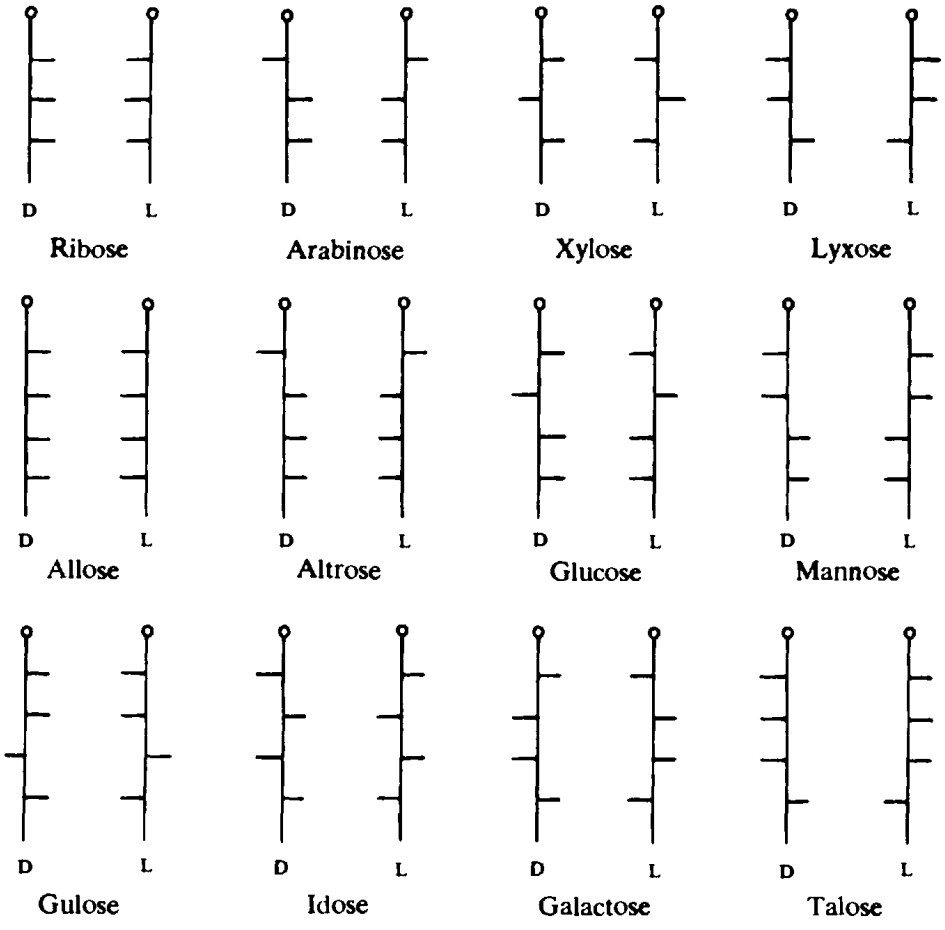
If the structures of the two tetroses are written alongside those of D-erythrose and D-threose using Fischer projection formulas, it is seen that two pairs of mirror images are given. That is, the four aldotetroses constitute two pairs of enantiomers:



Two simple aldopentoses can be derived structurally from each of the four aldotetroses described, making a total of eight aldopentoses. Therefore, there are 16 aldohexoses.

EXAMPLE 2.4

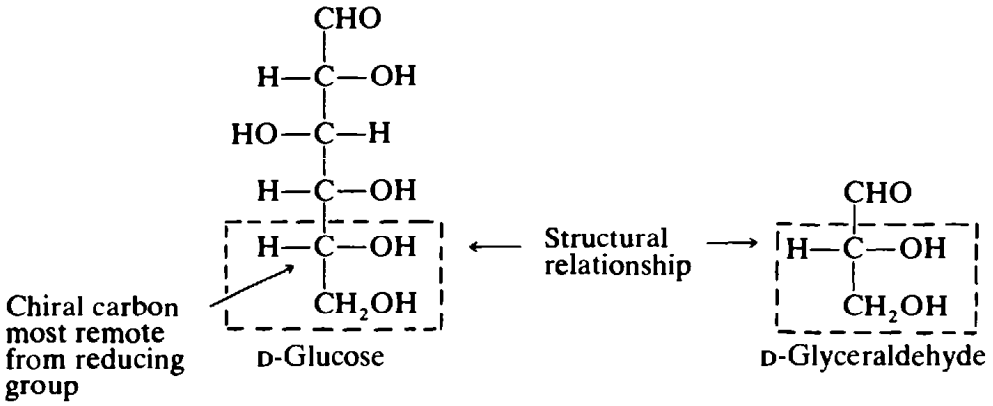
Simplified structures for the eight aldopentoses and 16 aldohexoses are shown (○ represents aldehyde; — represents an OH group; H atoms on carbons are omitted).



Question: Glyceraldehyde is sometimes known as *glycerose*. Why?

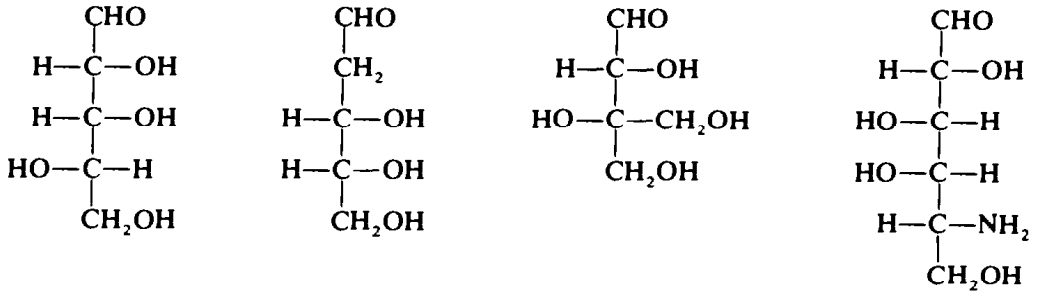
The names of all the aldotetroses, -pentoses, and -hexoses end in *-ose*. Glyceraldehyde is the parent aldose; thus, its name, glycerose, is valid.

There are two series of simple aldoses: a D series and an L series. To determine to which series an aldose belongs, locate the chiral carbon atom most remote from the reducing group and determine its relationship to glyceraldehyde; e.g., the sugar shown below, glucose, is called D-glucose:



EXAMPLE 2.5

In the sugars shown below, (1) is L and (2), (3), and (4) are D. Notice that in (3) the *chiral* carbon atom most remote from the reducing group is C-2.



Whereas glyceraldehyde has one chiral center, aldotetroses, -pentoses, and -hexoses have two, three, and four, respectively. Each chiral center gives rise to optical activity. The net optical activity of an aldose will depend on contributions from each chiral center and will be (+) or (−).

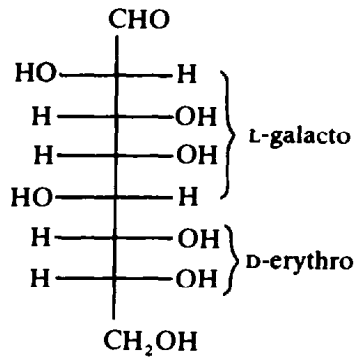
Question: When D-erythrose is dissolved in water, would you predict the solution to have a (+) or (−) optical rotation?

This is impossible to predict. The prefixes D- and L- refer to the *shape* of the molecule and imply nothing regarding the optical activity. In fact, a solution of D-erythrose is (−).

When more than four chiral carbon atoms are present, an aldose is given two configurational prefixes, one for the four lowest-numbered chiral centers and one for the rest of the molecule. The configuration of the *highest*-numbered group is stated first.

EXAMPLE 2.6

The aldooctose shown is named D-*erythro*-L-*galacto*octose.



2.4 SIMPLE KETOSES

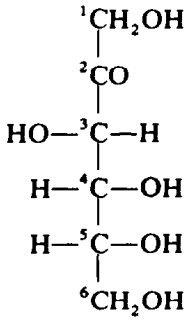
Structurally, the parent compound of the simple ketoses is *dihydroxyacetone*, a structural isomer of glyceraldehyde.

Dihydroxyacetone Glyceraldehyde

Although dihydroxyacetone does not possess a chiral carbon atom, the simple ketoses are related to it structurally by the introduction of hydroxylated chiral carbon atoms between the keto group and one of the hydroxymethyl groups. Thus there are two ketotetroses, four ketopentoses, and eight ketohexoses.

EXAMPLE 2.7

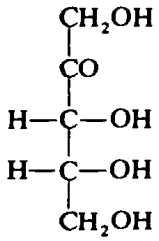
The most common ketose, D-*fructose*, is shown below. Compare the configuration of the chiral carbon atom most remote from the keto group (C-5) with D-glyceraldehyde.



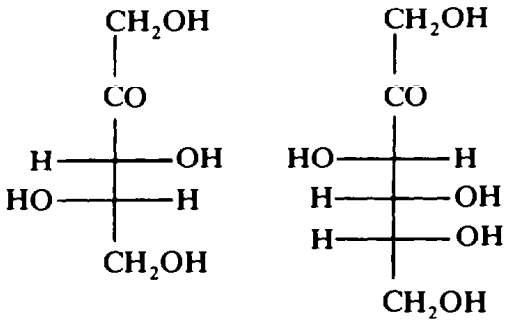
Fructose (shown in Example 2.7) was named long before its structure was known. The same is true for most aldoses. Names like glucose, mannose, ribose, and fructose are called *trivial names*; i.e., they are *nonsystematic*. However, ketoses, which are isomers of aldoses, were isolated or synthesized after the corresponding aldoses and were given names based on the names of the isomeric aldoses. Such names are misleading because they do not reflect the structural elements present in the ketoses.

EXAMPLE 2.8

The ketose shown below is known universally as D-*ribulose* because it is an isomer of D-ribose. The name is incorrect; the compound has only two chiral centers, *not* three as the prefix *rib-* would imply. The compound is related to D-erythrose, and its correct name is D-*erythro*-pentulose.



Question: What are the correct systematic names for the two sugars shown?



(1) (2)

(1) L-*Threo*-pentulose; commonly called L-xylulose
(2) D-*Arabino*-hexulose; commonly called D-fructose

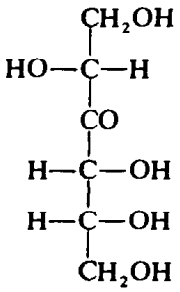
Question: How can you tell if a monosaccharide is a ketose from its name?

The systematic name always ends in *-ulose*, apart from fructose, and so do the trivial names.

Some ketoses are not related structurally to dihydroxyacetone. They are named by considering the configurations of all the chiral carbon atoms as a unit, ignoring the carbonyl group.

EXAMPLE 2.9

Consider the ketose shown. It has three chiral carbon atoms in the D-arabino configuration (even though interrupted by a keto group). The keto group is at position 3. The ketose has six carbon atoms. It is therefore called D-*arabino*-3-hexulose.



If the name of a ketose contains no number, it is assumed the ketose is related to dihydroxyacetone and the keto group is at position 2.

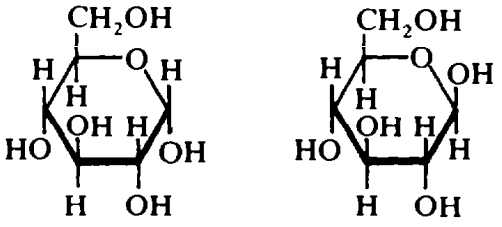
Question: Why is there no need to use numbers to indicate the position of the carbonyl group in aldoses?

The aldehyde group must always be terminal to a chain of carbon atoms.

2.5 THE STRUCTURE OF D-GLUCOSE

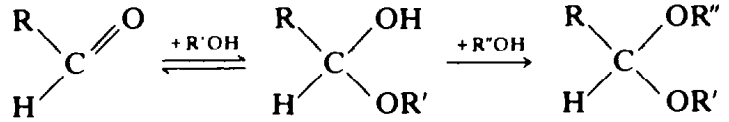
D-Glucose is the most common of the monosaccharides, occurring in the free state in the blood of animals and in the polymerized state, inter alia, as starch and cellulose. Tens of millions of tons of these polysaccharides are made by plants and photosynthetic microbes annually. A detailed study of the structure of glucose is justified on these grounds, and many of the structural features of all monosaccharides can be illustrated using glucose as an example.

The Fischer projection formula for D-glucose (Example 2.3) is also known as the *open-* or *straight-*chain structure. This structure occurs only in solution. There are two crystalline forms of D-glucose, known as α and β, which also have different optical activities when dissolved. X-ray diffraction studies have confirmed chemical evidence that α- and β-D-glucose are structures containing a ring of five carbon atoms and one oxygen atom:

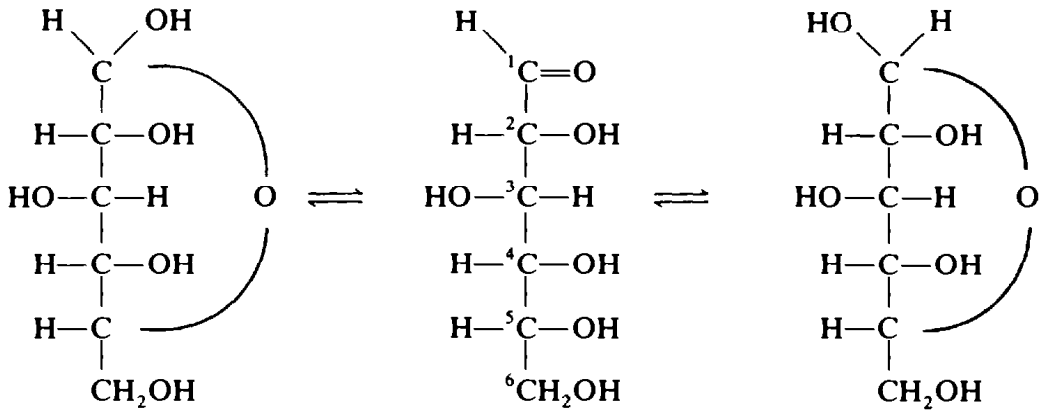


These structures are known as the *Haworth projection formulas*. They do not represent the true shapes of the molecules (Sec. 2.6), but they do show the configuration at each of the chiral atoms. The formulas are meant to represent the ring as a generalized plane standing at right angles to the plane of the paper and an element of perspective is given to the structure by thickening the three bonds of the ring that are meant to appear to be in front of the paper; the remaining three bonds of the ring and the oxygen atom in the ring lie behind the paper.

When α-D-glucose or β-D-glucose is dissolved in water, the ring opens and the open-chain structure is formed. The reaction is *reversible*, and an equilibrium is established between the open form and the two ring forms. The chemistry of the process is understood in terms of the chemistry of the aldehyde group. In general, aldehydes react reversibly with alcohols to give *hemiacetals* and then, in the presence of an acid catalyst (Chap. 8), *acetals*:



The formation of a ring by the open-chain form of D-glucose can be considered to be the result of a reaction between the hydroxyl group on C-5 and the aldehyde group to give a hemiacetal. The aldehyde carbon becomes chiral as a result, thus giving rise to two hemiacetals, α- and β-D-glucose.



A hemiacetal (α-D-glucose) A hemiacetal (β-D-glucose)

These hemiacetals are known as *anomers*, and C-1 of the ring form is called the *anomeric* carbon.

Question: Why is the anomeric carbon sometimes called the *potential reducing carbon*?

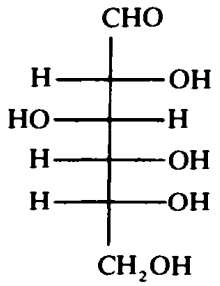
When the ring opens in solution, C-1 becomes the carbon atom of the aldehyde group, which has reducing properties. Only the open-chain form has reducing properties; the ring forms are nonreducing because they lack an aldehyde group.

The Haworth projection formulas are neater ways of writing the ring forms shown in the equilibria above and yet preserving the configuration shown at each chiral carbon. It is not difficult to translate the open-chain structure for a monosaccharide into the Haworth ring structure.

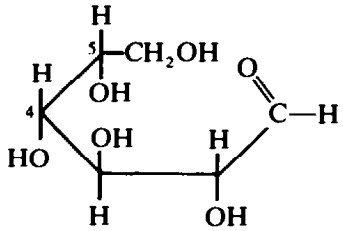
EXAMPLE 2.10

The translation of the open-chain structure into the Haworth ring structure is best achieved in a step-by-step procedure, since a direct translation is tricky. D-Glucose is used in this example:

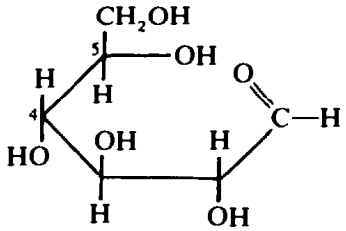
1. Write the open-chain structure.



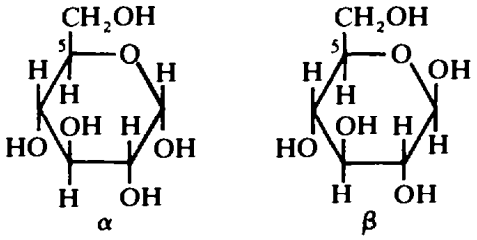
2. Turn the structure clockwise on its side and bend it round to almost form a ring.



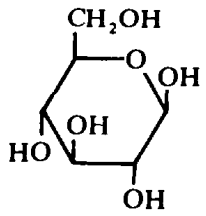
3. Rotate the C-4—C-5 bond to bring the hydroxyl on C-5 close to the carbonyl group.



4. Form the hemiacetal(s) by bonding of the hydroxyl on C-5 to the carbonyl group.



Note: It is normal in writing the ring structures of sugars to omit the H atoms attached to carbon, in which case β-D-glucose is written as:



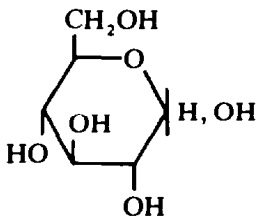
Question: What determines which of the anomers shown in Example 2.10 is called α and which β?

This depends on the configuration of the anomeric carbon relative to the configuration of the chiral atom that establishes whether the monosaccharide belongs to the L or the D series (in the case of glucose, C-5). If, in the Fischer projection formula, the hydroxyl groups on these carbons are cis, the anomer is called α; if the hydroxyl groups are trans, the anomer is called β. A comparison of the structure in step (2) of Example 2.10 with those in step (4) shows that the hydroxyl group on C-5 was below the plane of the ring before the rings were closed. Thus, the anomer with the anomeric hydroxyl below the plane of the ring is called α.

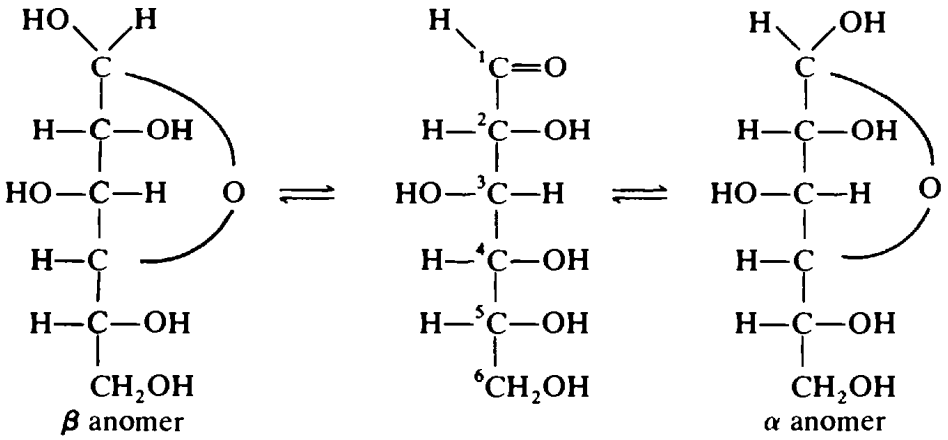
Although the open-chain form of glucose has four chiral centers, formation of the ring creates a fifth chiral center, which is, of course, why there are two anomers that differ in optical rotation. When solid α-D-glucose is dissolved in water, $[\alpha]_D^{25}$ is +112°. When solid β-D-glucose is dissolved in water, $[\alpha]_D^{25}$ is +19°.

EXAMPLE 2.11

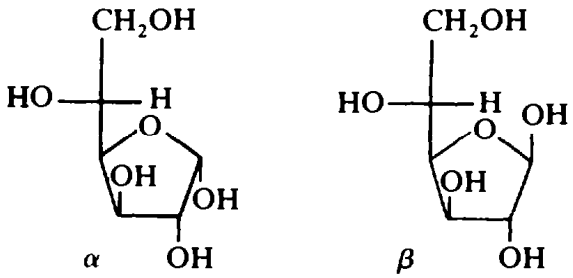
The optical rotations of freshly made solutions of α- and β-D-glucose change with time. This is called *mutarotation*. A change in optical rotation must mean a change in structure. When α-D-glucose is dissolved in water, the α-D anomer is the only structure present at the instant of dissolution. The opening of the ring is a slow reaction, but since the reaction is reversible, some β-D anomer as well as the open-chain form will appear. Ultimately, an equilibrium between the open-chain form and the two anomers is established with an $[\alpha]_D^{25}$ of +52°. For this reason, the structure shown below is used to describe the state of glucose in solution, and the same device is used to depict the structure of a sugar in which the configuration of the anomeric carbon is unknown.



Although we have considered the anomeric forms of D-glucose to consist of six-membered rings, it is possible for two other anomeric forms to arise from the open-chain form by the addition of the hydroxyl group on C-4 to the carbonyl group of the aldehyde. These are five-membered ring forms.



Question: How do we translate the open-chain form of glucose into five-membered Haworth projection formulas and assign the symbols α and β to the appropriate formula?



Clearly, it is necessary to distinguish between six- and five-membered rings. This is done by expanding the name *glucose* to *glucopyranose* for the six-membered anomers and to *glucofuranose* for the five-membered anomers.

Question: Would you expect furanose or pyranose ring forms to be the more stable for a given sugar?

Furanoses are *generally* less stable because the area of the ring is smaller and the opportunity for destabilizing the structure through steric interference between the H, hydroxyl, and hydroxymethyl groups on different carbons is greater. With glucose, the α- and β-glucofuranoses contribute very little to the equilibrium mixture.

2.6 THE CONFORMATION OF GLUCOSE

It was stressed in the previous section that the Haworth structures for the anomers of D-glucopyranose do not represent the true shape of the rings. The carbon atoms of glucose are all saturated, and the most stable form of a ring will be one that is *strain free*, i.e., where the angles formed by the bonds at each carbon atom are 109°, the tetrahedral angle.

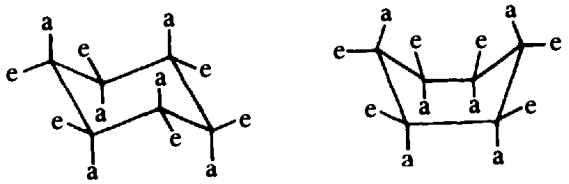
EXAMPLE 2.12

For simplicity, consider cyclohexane. Only two strain-free conformations are possible: a *chair* and a *boat*, which, for clarity, are shown without H atoms:

C_6H_{12}

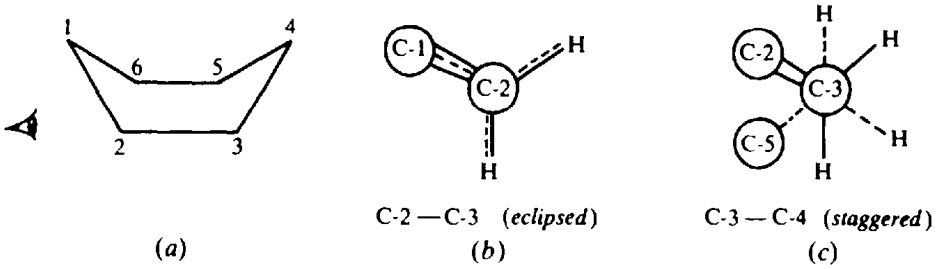
Cyclohexane Chair Boat

There are 12 H atoms in the structure; 6 are in the general plane of the rings and are called *equatorial* (e); 6 are perpendicular to the general plane of the rings and are called *axial* (a).



The chair and boat forms are interconvertible by rotation around the C—C bonds. Therefore, they are not isomers, and the term *conformers* is used to describe the various shapes a molecule can possess.

Although both conformers of cyclohexane are strain-free, the chair conformer is more stable than the boat. There are two reasons for this. (1) Two of the six C—C bonds in the boat are *eclipsed*, whereas none are in the chair conformation. This can be illustrated by imagining the eye placed to look along the C-2—C-3 bond, as shown in Fig. 2-1. The eye, positioned as in (*a*), will see the configuration shown in (*b*): C-2 with its two H atoms and the bond leading to C-1; however, the eye cannot see C-3 or the two H atoms attached to it, because these three atoms are immediately behind C-2 and its two H atoms. Likewise, the C-5—C-6 bond is eclipsed, but all other C—C bonds in the boat and *all* C—C bonds in the chair are *staggered*, e.g., (*c*) shows the eye looking along C-3—C-4. Thus in the boat, the four axial H atoms on C-2, C-3, C-5, and C-6 are as close as possible, and interactions between them tending to distort the ring are maximal. (2) The axial H atoms on C-1 and C-4 of the boat come closer together (0.18 nm) than the sum of their van der Waals radii (0.24 nm), and with these two atoms there is an unfavorable interaction tending to distort the ring.



C-2 — C-3 (*eclipsed*) C-3 — C-4 (*staggered*)

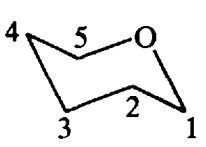
(*a*) (*b*) (*c*)

Fig. 2-1

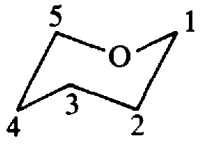
The difference in energy between the boat and chair conformers of cyclohexane is about 25 kJ mol^{-1}, which means that at 25°C only 1 in 1,000 molecules exists in the boat conformation.

Question: What are two ways in which a monosaccharide in a ring form, like glucose, differs from cyclohexane in conformation?

(1) Monosaccharides have hydroxyl and hydroxymethyl groups replacing some of the H atoms on the carbon atoms of the ring. These are much more bulky than H atoms and would tend to be in equatorial positions around the edge of the ring rather than in axial positions on the faces of the ring, where they would be closer together and thus distort the ring. (2) Monosaccharides have an oxygen atom in the ring in the place of a carbon. This has little effect on the shape of the ring since, although the C—O bond length is slightly shorter than the C—C bond length, the valencies of oxygen are at 109°. However, the presence of an oxygen atom in the ring does mean there are *two* possible chair conformers, known as *C1* and *1C*.



C1



1C

These cannot be superimposed, but C1 can be converted reversibly to 1C by rotation of the bonds, passing through a boat conformer in the process.

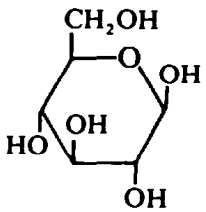
A Haworth structure for a monosaccharide is translated readily into a structure showing the true shape of the molecule.

EXAMPLE 2.13

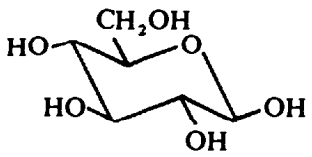
The Haworth structure tells us that a substituent that is above the general plane of the monosaccharide ring must also appear above the general plane of the chair; whether this is axial or equatorial will depend on the carbon atom being considered. For example, looking at each carbon of β-D-glucopyranose in turn, we find

Substituent	Haworth Structure—Position Relative to Ring	Chair Conformers*	
		C1	1C
OH of C-1	above	e	a
H of C-1	below	a	e
OH of C-2	below	e	a
H of C-2	above	a	e
OH of C-3	above	e	a
H of C-3	below	a	e
OH of C-4	below	e	a
H of C-4	above	a	e
CH_2OH of C-5	above	e	a
H of C-5	below	a	e

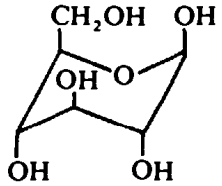
*e = equatorial plane; a = axial plane.



Haworth



C1

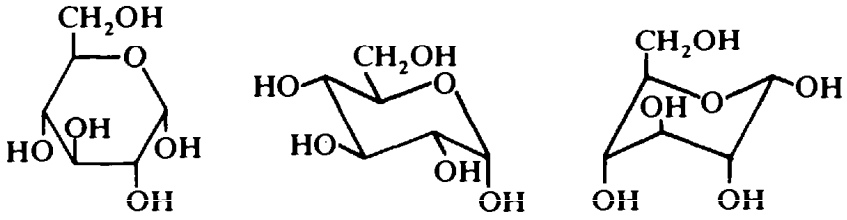


1C

The preferred conformer is C1, where all the bulky substituents are equatorial, on the edge of the ring. There will be considerable distortion in the 1C conformer, where three bulky groups on the upper face of the ring produce considerable steric hindrance, as do two bulky groups on the lower face.

Question: Which is the preferred conformer for α-D-glucopyranose?

Proceeding as in Example 2.13, we see that C1 is again the preferred conformer:



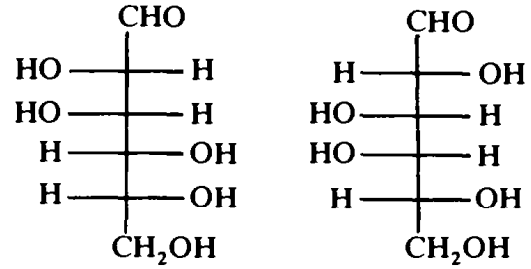
C1 1C

If the preferred forms of α- and β-D-glucopyranose (both C1) are compared, it is seen that the β form is the more stable since it has no bulky axial substituent, whereas the α form has one. This helps explain why β-D-glucopyranose is the dominant anomer in an aqueous solution of glucose.

2.7 MONOSACCHARIDES OTHER THAN GLUCOSE

Aldohexoses

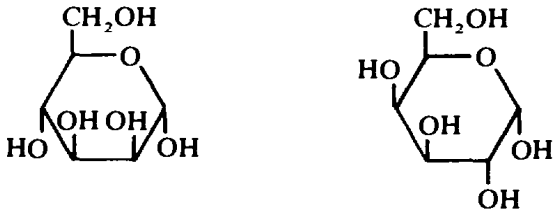
Two structural isomers of glucose are mannose and galactose.



D-Mannose D-Galactose

EXAMPLE 2.14

D-Mannose and D-galactose are readily translated into six-membered Haworth structures. The α anomers only are shown.

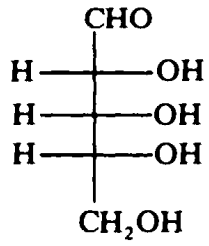


α-D-Mannopyranose α-D-Galactopyranose

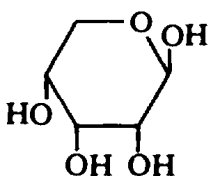
Both of these are *epimers* of glucose; that is, neglecting the anomeric carbon, they differ from glucose in the configuration at just one carbon atom—mannose at C-2, galactose at C-4.

Aldopentoses

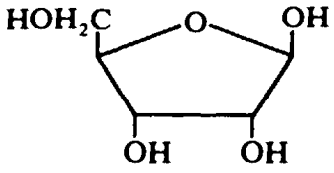
The most common aldopentose is D-ribose, the sugar present in RNA.



Question: Translate D-ribose into five- and six-membered Haworth structures. Give the β anomers only.



β-D-Ribopyranose



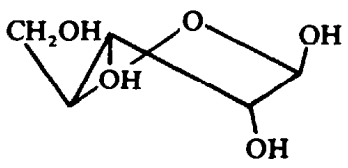
β-D-Ribofuranose

The only known *crystalline* form of D-ribose is β-D-ribopyranose, but in solution the α and β anomers of both pyranose and furanose forms occur.

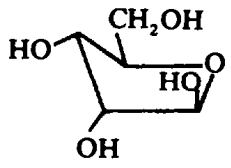
The angle of a regular pentagon is 108°, very close to the tetrahedral angle, suggesting that the shape of a furanose ring is nearly planar. With some sugars having a furanose ring, the structure cannot be planar because of steric repulsion between substituents on the ring.

EXAMPLE 2.15

In β-D-ribofuranose, notice that the hydroxyl groups and H atoms attached to the carbon atoms of the C-2—C-3 bond are eclipsed. Repulsions between these eclipsed groups would be relieved by twisting the ring out of the plane, i.e., by raising C-3 above the plane and dropping C-2 below the plane. This conformer is known as T_2^3 (T for *twist*, with C-3 raised and C-2 lowered). An alternative way of relieving steric repulsion is to raise or lower C-3 out of the plane of the ring. This gives E (for *envelope*) conformers.



T_2^3

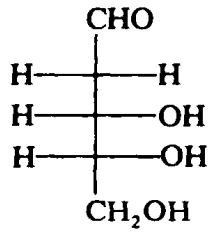


E^3

The loss of stability of the ring when C-3 is moved out of the plane is more than compensated by the relief from steric repulsion.

Deoxy Sugars

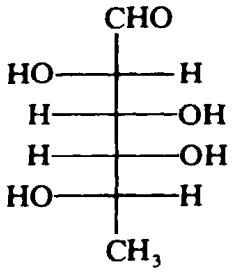
Deoxy sugars are reduced forms of sugars in which a hydroxyl group is replaced by a hydrogen atom. The most widely distributed deoxy sugar is known as 2-deoxy-D-ribose and is present in DNA.



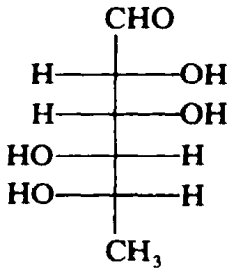
EXAMPLE 2.16

As with ketoses, many deoxy sugars are incorrectly named. In so-called 2-deoxy-D-ribose, there are only two chiral carbon atoms and the sugar is related to D-erythrose. The systematic name is 2-deoxy-D-*erythro*-pentose.

The other deoxy sugars commonly found are L-fucose, particularly in animals, and L-rhamnose, occurring in plants and bacteria.



L-Fucose



L-Rhamnose

Question: What are the systematic names for these sugars?

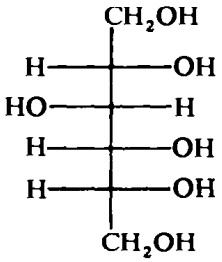
6-Deoxy-L-galactose (L-fucose) and 6-deoxy-L-mannose (L-rhamnose).

Alditols

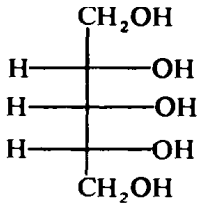
A different type of reduced sugar is an *alditol*, in which the aldehyde group of an aldose has been reduced. For example, the alditol produced from D-glucose is D-glucitol (the trivial name is sorbitol). The name of an alditol is obtained by adding *-itol* to the root of the name of the aldose (except for glycerol, a reduction product of glyceraldehyde).

EXAMPLE 2.17

Two alditols are shown below.

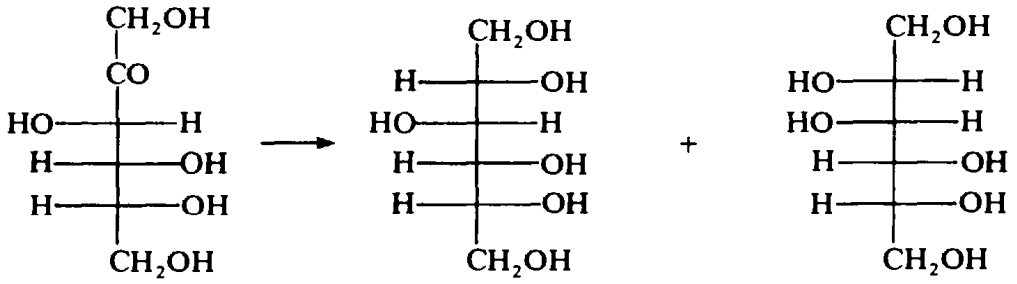


D-Glucitol



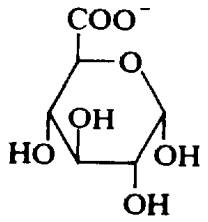
D-Ribitol

Question: What are the names of the compounds obtained when the carbonyl group of D-fructose is reduced to an alcohol?



Uronic and Aldonic Acids

Uronic acids are sugars in which the hydroxymethyl group of an aldose has been oxidized to a carboxylic acid. These occur as the salts, known as *uronates*, at physiological pH.

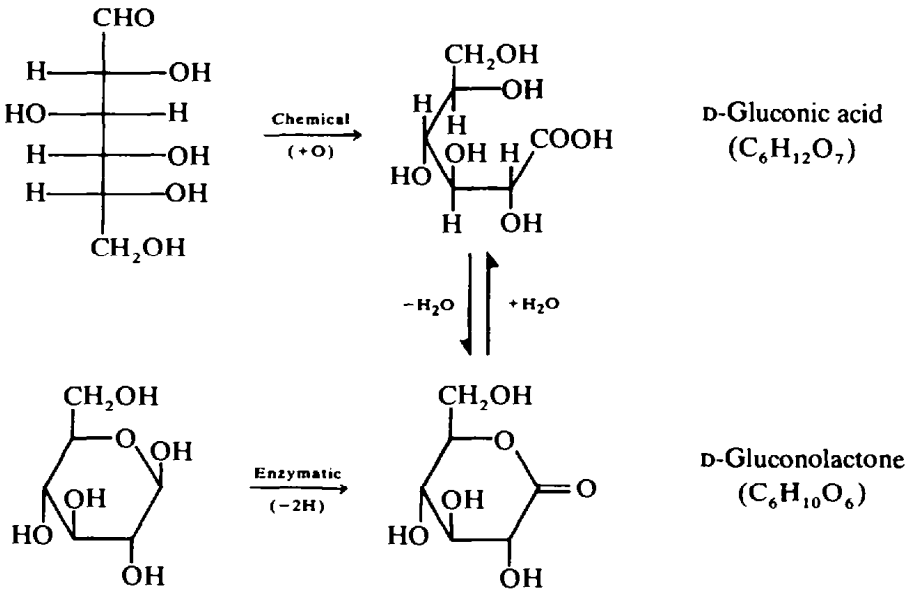


α-D-Glucuronate

Oxidation of the aldehyde group of an aldose to a carboxylic acid group gives a derivative known as an *aldonic acid*. This occurs as a salt, *aldonate*, at physiological pH. If an aldonic acid contains five or more carbon atoms, a *δ-lactone* is formed spontaneously by the condensation of the carboxylic acid group and the hydroxyl group on C-5.

EXAMPLE 2.18

With glucose, chemical oxidation of C-1 occurs with the open-chain form, whereas enzyme-catalyzed oxidation of C-1 occurs with the ring form of the sugar. The aldonic acid (gluconic acid) that is formed is an equilibrium mixture of the free acid (open chain) and the δ-lactone (ring).



Amino Sugars

Amino sugars are widely distributed naturally. Generally, they are sugars in which a hydroxyl group has been replaced by an amino group.

EXAMPLE 2.19

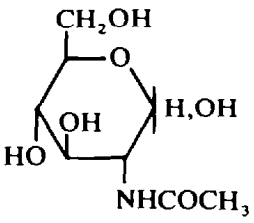
The most common amino sugars are:

D-Glucosamine (2-amino-2-deoxy-D-glucose)

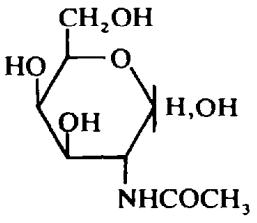
D-Galactosamine (2-amino-2-deoxy-D-galactose)

Neuraminic acid, a derivative of D-mannosamine (2-amino-2-deoxy-D-mannose)

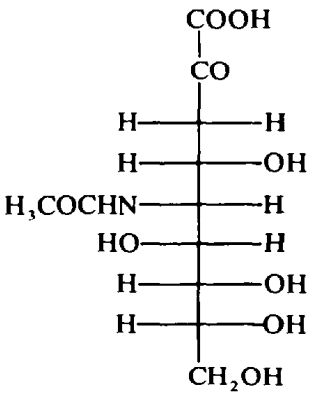
The amino groups in these compounds are usually acetylated.



N-Acetyl-D-glucosamine



N-Acetyl-D-galactosamine



N-Acetylneuraminic acid (sialic acid)

Phosphate and Sulfate Esters

Many monosaccharides and their derivatives occur naturally in a form in which one or more of the hydroxyl groups has been substituted by a phosphate or a sulfate group. These are known as *esters*. In general, the phosphate esters are found as components of metabolic pathways within cells, whereas the sulfate esters are found in oligosaccharides and polysaccharides occurring outside cells.

EXAMPLE 2.20

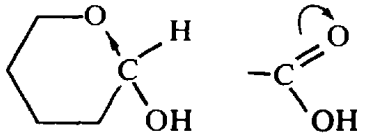
Two phosphate esters and two sulfate esters of monosaccharides that occur naturally are fructose 1,6-bisphosphate and 6-phosphogluconate (for both, see Chap. 11); and D-galactose 4-sulfate and *N*-acetylgalactosamine 4-sulfate.

2.8 THE GLYCOSIDIC BOND

All monosaccharides and their derivatives that possess aldehyde or ketone groups (that is, excepting derivatives such as alditols and aldonic acids) will have reducing properties. Moreover, those with the appropriate number of carbon atoms can form rings occurring in two forms (anomers) and in which the potential reducing carbon is called the anomeric carbon.

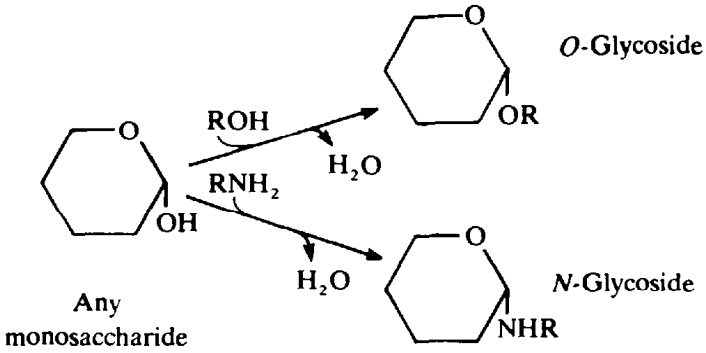
Question: How reactive is the anomeric hydroxyl group compared with the other hydroxyl groups in a monosaccharide?

It is much more reactive than a typical primary or secondary alcohol. This reactivity is due to the electron-withdrawing influence of the ring oxygen. The situation may be compared to the structure written for a carboxylic acid in which the four atoms, C-O and O-H, constitute the chemical group. With a monosaccharide the anomeric C atom, the H and O-H atoms on the anomeric carbon, and the ring oxygen constitute the corresponding chemical group.



EXAMPLE 2.21

The reactivity of the anomeric hydroxyl group is illustrated by the ease with which monosaccharides react with alcohols and with amines. The normal hydroxyl groups in the molecule do not react, but the anomeric hydroxyl does. The process is known as *glycosylation* (of the alcohol or amine), and the products as O-*glycosides* and N-*glycosides*.



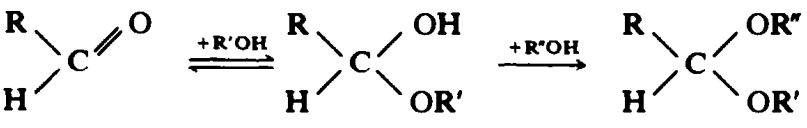
The R group in these glycosides is referred to as the *aglycone*.

As with the parent monosaccharide, the anomeric carbon can have either the α or the β configuration since it remains a chiral center.

Question: In what respect does a glycoside ring differ from the ring structure of a monosaccharide?

The ring of a glycoside cannot open to give a straight-chain structure. Consequently, glycosides have no reducing properties.

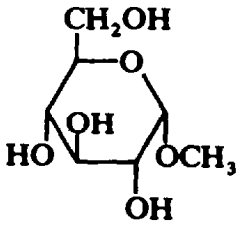
A glycoside is, in chemical terms, an acetal (see also Sec. 2.5):



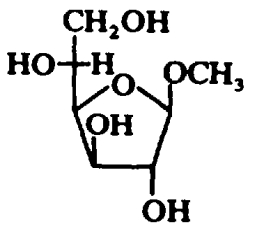
Aldehyde (open-chain monosaccharide) **Hemiacetal (ring form)** **Acetal (glycoside)**

The term *glycoside* is a generic one, and glycosides derived from glucose, fructose, and ribose are known as *glucosides*, *fructosides*, *ribosides*, etc.; i.e., *-oside* is the suffix for glycosides. Glycosides,

like the parent monosaccharides, can have either five- or six-membered rings known as *furanosides* or *pyranosides*, respectively.



α-D-Methylglucopyranoside



β-D-Methylglucofuranoside

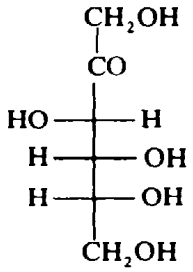
Question: What happens to the optical rotation when a freshly prepared solution of α-D-methylglucopyranoside is allowed to stand?

It remains constant. Since a glycosidic ring cannot open, these structures do not display mutarotation.

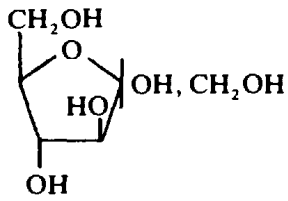
EXAMPLE 2.22

Give the structure for β-D-methylfructofuranoside.

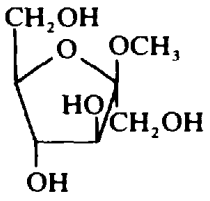
1. Write the structure for D-fructose.
2. Convert this to the five-membered Haworth ring form (use Example 2.10 as a guide).
3. Take the β anomer and substitute the anomeric hydroxyl with a methyl group.



(1)



(2)



(3)

β-D-Methylfructofuranoside

The glycosidic bond is found in a very wide range of biological compounds. In addition, there is a particularly important group of *O*-glycosides in which the glycosidic bond links two monosaccharides, i.e., the aglycone is a sugar. Such compounds are called *disaccharides*. The anomeric hydroxyl group of the second monosaccharide can itself glycosylate a hydroxyl group in a third monosaccharide to give a trisaccharide, and so on. *Polysaccharides* are polymers in which a large number of monosaccharides are linked by glycosidic bonds.

Question: What is an oligosaccharide?

An *oligosaccharide* is a compound consisting of an undefined but small number of monosaccharides linked by glycosidic bonds.

Question: Why are the monosaccharide units in oligo- and polysaccharides called *residues*?

The formation of each glycosidic bond is a *condensation reaction* in which a water molecule is produced. Effectively then, all the monosaccharides, except the monosaccharide at one end, have lost a water molecule; thus the term *residue* is justified.

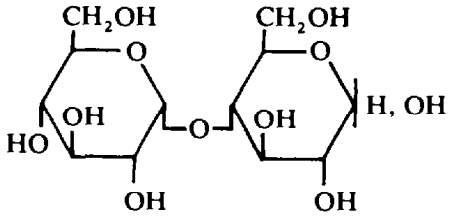
The name of a residue is formed by adding *-osyl* to the root of the name of the sugar. Thus, a trisaccharide made from three glucose molecules is glucosylglucosylglucose. In order to abbreviate the written descriptions for oligo- and polysaccharides, a shorthand method for naming residues has been introduced; e.g., Glc (glucosyl), Gal (galactosyl), Fru (fructosyl), GlcN (glucosaminyl), GlcNAc (*N*-acetylglucosaminyl), GlcA (glucuronyl), NeuNAc (*N*-acetylneuraminyl).

The term *glycan* is used as an alternative to the word *polysaccharide*. This is a generic term, and names such as *glucan*, *xylan*, *glucomannan* describe polymers composed, respectively, of glucose residues, xylose residues, and glucose and mannose residues.

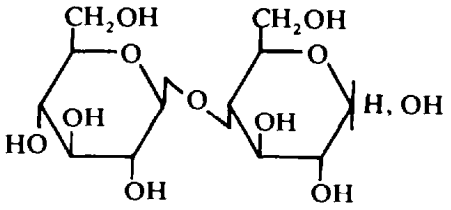
The greatest problem in understanding the structures of oligo- and polysaccharides is related to the ways in which the glycosidic bonds are written.

EXAMPLE 2.23

The structures of two different disaccharides, both composed of glucose, are shown.

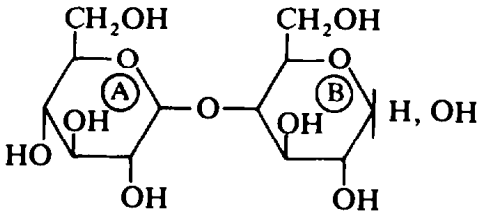


Maltose
α-Glc-(1→4)-Glc



Cellobiose
β-Glc-(1→4)-Glc

The glycosidic bonds joining the glucose molecules are printed the way they are for the following reason. Consider a disaccharide formed between two monosaccharides, A and B, in which the anomeric hydroxyl of A is used to glycosylate the hydroxyl on C-4 of B:



This structure is ambiguous in two respects. It does not show the configuration at C-1 of A, i.e., whether the glycosidic bond is α or β. Nor does it show the configuration at C-4 of B, and therefore the identity of B is not revealed. B could be either glucose (with the glycosidic O below the ring B) or galactose (with the glycosidic O above ring B). Thus maltose and cellobiose are written as shown to indicate clearly (1) whether the glycosidic bond is α or β with respect to the glycosyl component and (2) what the identity of the other sugar is. The printing of the glycosidic bonds as C—O—C and C⌒O⌒C is a pictorial device: it gives clearly the correct configuration of the chiral carbon atoms. The bonds in reality are not bent.

Question: What are the systematic names for maltose and cellobiose?

Maltose is α-D-glucosyl-(1→4)-D-glucose or, more specifically, α-D-glucopyranosyl-(1→4)-D-glucopyranose. Cellobiose is β-D-glucosyl-(1→4)-D-glucose.

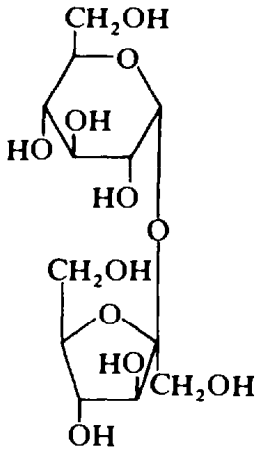
In the structure for maltose shown in Example 2.23, no configuration is given for the anomeric carbon of the glucose unit on the right; the structure represents the state of maltose in solution—a mixture of α and β anomers. In crystalline maltose, the anomeric hydroxyl is α, and maltose can be described as α-D-glucosyl-(1 → 4)-α-D-glucose.

Question: Is maltose a reducing sugar?

Yes. Although it is a glycoside, the second glucose unit possesses an anomeric carbon atom and its ring can open to give an aldehyde. For the same reason, solutions of maltose display mutarotation.

Question: Why doesn't sucrose, another disaccharide, have reducing properties?

Sucrose, or cane sugar, is a disaccharide in which the anomeric hydroxyl of α-D-glucose is condensed with the anomeric hydroxyl of β-D-fructose (Example 2.22). It is therefore both an α-glucoside and a β-fructoside. Neither unit possesses an anomeric hydroxyl and neither ring can open to give an aldehyde.



Question: What is the systematic name for sucrose?

Either α-D-glucosyl-(1 → 2)-β-D-fructoside (α-D-glucopyranosyl-(1 → 2)-β-D-fructofuranoside) or β-D-fructosyl-(2 → 1)-α-D-glucoside (β-D-fructofuranosyl-(2 → 1)-α-D-glucopyranoside). In abbreviated form, these are α-Glc-(1 → 2)-β-Fru and β-Fru-(2 → 1)-α-Glc, respectively.

2.9 POLYSACCHARIDES

Polysaccharides function mainly either as structural components or as forms of energy storage.

EXAMPLE 2.24

Starch and *glycogen* are polymers of glucose (or glucans) that serve as sugar storage molecules in plants and animals, respectively. Starch exists as *amylose*, a linear polymer of α-D-glucose joined with α(1 → 4) linkages, and as *amylopectin*, in which branching occurs through additional α(1 → 6) linkages. Glycogen (Chap. 11) is also a branched glucan, but the degree of branching is greater in glycogen than in amylopectin. The long linear chains of amylose can exist with adjacent residues randomly oriented with respect to each other, although there is a tendency for the chains to take up a helical conformation. In the presence of iodine, interactions between the glucose residues and iodine atoms stabilize the helical conformation, with iodine atoms enclosed within the helix, and results in the development of a structure that has a bright blue color.

EXAMPLE 2.25

Cellulose, the major cell wall material of plants, is also a linear polymer of glucose, but connected by $\beta(1 \rightarrow 4)$ linkages.

Question: Why is cellulose insoluble, while starch, which appears to have a very similar structure, is soluble?

The seemingly small difference in structure between starch and cellulose allows the linear chains of cellulose to pack together side-by-side in an antiparallel extended conformation, stabilized by hydrogen bonds, to produce an insoluble structure of high mechanical strength.

The $\beta(1 \rightarrow 4)$ linkage is particularly stable with respect to hydrolysis. Cellulose cannot be digested by mammals, but some insects (notably termites and wood-eating cockroaches), protozoans and fungi possess *cellulases*, enzymes that can hydrolyze the $\beta(1 \rightarrow 4)$ linkages. Ruminants, such as sheep and cattle, can digest cellulose because of the protozoans that live symbiotically in their digestive system.

In addition to cellulose, plants contain *pectins* and *hemicelluloses*. Pectins are polymers comprised of arabinose, galactose and galacturonic acid. Hemicelluloses are not derived from cellulose, but are polymers of D-xylose, D-mannose or D-galactose. These compounds provide a matrix in which the cellulose fibers are embedded.

The cell walls of many fungi, and the exoskeletons of insects and other arthropods, as well as some molluscs, are largely comprised of *chitin*, a polymer of *N*-acetyl-β-D-glucosamine linked by $\beta(1 \rightarrow 4)$ glycosidic bonds.

The connective tissue of mammals is rich in polysaccharides known as *glycosaminoglycans*, or *mucopolysaccharides*. Chondroitin sulfate is a polymer of a repeating unit comprising $\alpha(1 \rightarrow 3)$ linked disaccharide of glucuronic acid and *N*-acetylgalactosamine; the galactosamine is largely sulfated on the hydroxyl of carbon 6. Hyaluronic acid, a lubricating material in the synovial fluid of joints such as the elbow and knee, is a repeating polymer of a disaccharide of glucuronic acid and *N*-acetylglucosamine.

Question: What is the (biological) function of the carboxyl and sulfate groups on glycosaminoglycans?

The negative charges on these groups at neutral pH lead to electrostatic repulsion which increases the dimensions of the molecule. The large volume of solution occupied by such expanded molecules contributes to high viscosity, and aids in the function of hyaluronic acid as a joint lubricant.

Solved Problems

THE STRUCTURE OF D-GLUCOSE

2.1. A solution of D-glucose contains predominantly the α and β anomers of D-glucopyranose, both of which are nonreducing. Why is a solution of D-glucose a strong reducing agent?

SOLUTION

Because there is some open-chain glucose present with reducing properties. As this reacts, the equilibria between it and the nonreducing ring forms are disturbed, causing more of the open-chain form to appear. Ultimately, all the glucose will have reacted via the open-chain form.

2.2. What percentage of D-glucose in solution at equilibrium exists as the β anomer?

SOLUTION

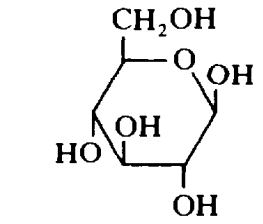
By definition, 1 g of β-D-glucose in 1 cm^3 will give a rotation of $+19°$ in a 1-dm polarimeter tube. Likewise 1 g of α-D-glucose will give a rotation of $+112°$ (see Sec. 2.5). At equilibrium, let there be b g cm^{-3} of β-D-glucose. There will be $(1-b)$ g cm^{-3} of α-D-glucose at equilibrium. The rotation at equilibrium ($+52°$; Example 2.11) will be due to a contribution from b g of β-D-glucose and $(1-b)$ g of α-D-glucose; i.e., $b(+19°) + (1-b)(+112°) = +52°$, from which $b = 0.64$ g. Therefore the percentage of glucose present as the β anomer at equilibrium is 64 percent. (This answer assumes that the open-chain form of glucose makes a negligible contribution to the equilibrium mixture, an assumption supported by optical and NMR spectroscopy of the solution at equilibrium.)

THE CONFORMATION OF GLUCOSE

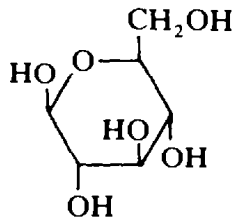
2.3. Write the Haworth structure for β-L-glucopyranose.

SOLUTION

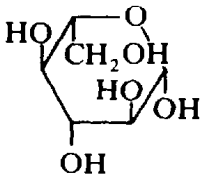
By definition, this is the mirror image of β-D-glucopyranose. The structure can be drawn in six ways by imagining a mirror placed (1) to the left, (2) to the right, (3) above, (4) below, (5) in front of, or (6) behind the structure of β-D-glucopyranose. Three of these images are shown.



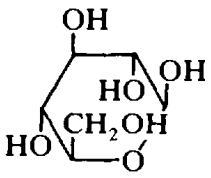
β-D-Glucopyranose



(2)



(4)



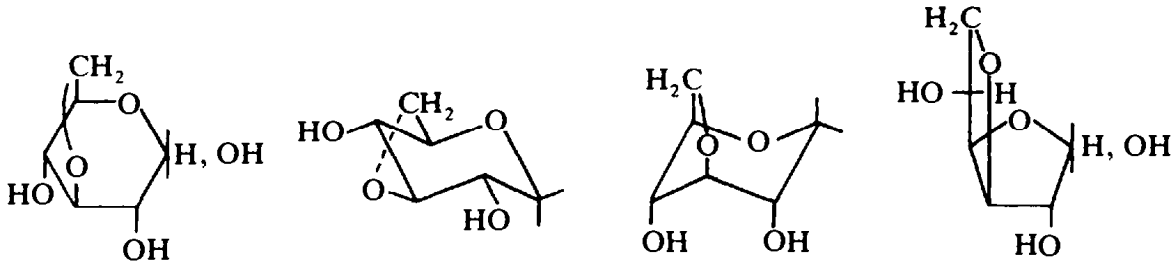
(5)

All six images represent the same structure and can be seen to be so by rotating any image about an appropriate axis; e.g., rotating (4) toward you through 180° about a horizontal axis in the plane of the paper will give (5).

2.4. What is the preferred conformation for 3,6-anhydro-D-glucose?

SOLUTION

3,6-Anhydro-D-glucose (1) is a glucose molecule in which a water molecule is lost between C-3 and C-6. Although the preferred conformation of D-glucose is C1, this conformation is not possible with (1) because the oxygen on C-3 is too far from C-6 for a bond to be formed between them (2). Consequently, 1C is a more likely conformation (3), but the large number of axial hydroxyl groups and the strain in the ring bounded by C-3, C-4, C-5, C-6 and the O between C-3 and C-6 make this conformation unstable. The preferred conformation for 3,6-anhydro-D-glucose is (4), with two fused, planar five-membered rings.



(1) (2) (3) (4)

MONOSACCHARIDES OTHER THAN GLUCOSE

2.5. Are D-mannose and D-galactose epimers?

SOLUTION

No. They differ in configuration at *two* carbon atoms.

2.6. Figure 2-2 shows the mutarotation that occurs when, β-D-ribose is dissolved in water. How can this curve be interpreted?

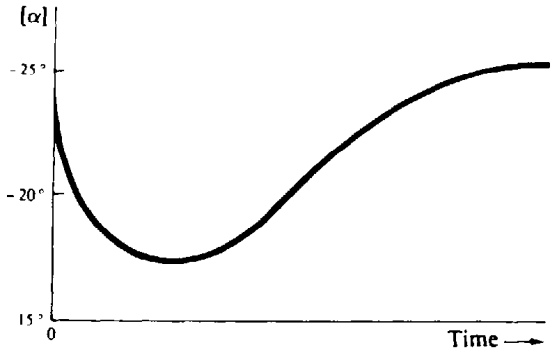


Fig. 2-2 Mutarotation of β-D-ribose dissolved in H_2O.

SOLUTION

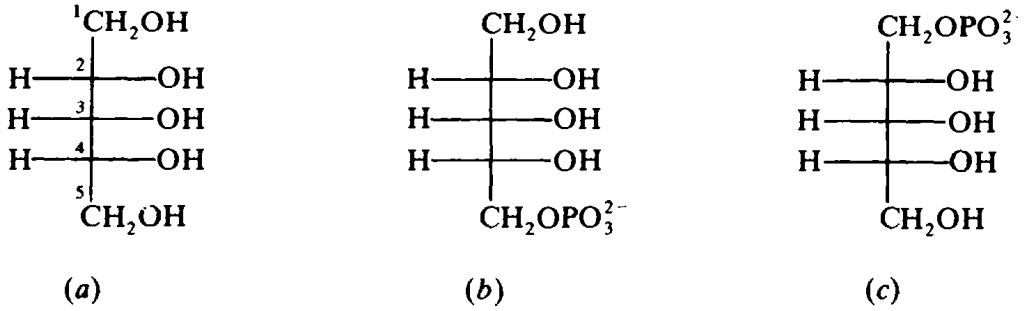
Crystalline β-D-ribose is β-D-ribopyranose (Sec. 2.7), and when this ribose is dissolved in water, the pyranose ring opens. The resulting open-chain form will close to give furanose and pyranose ring forms. Formation of the furanose ring is much faster since the random motion of the open-chain form allows the hydroxyl group on C-4 to approach C-1 far more frequently than the hydroxyl group on C-5 can. However, the pyranose ring is more stable because it is larger and the steric repulsion is less. Thus, the initial changes in rotation shown in Fig. 2-2 are due to the appearance of a relatively high concentration of furanose forms, and the later changes in rotation reflect the reappearance of increasing concentrations of pyranose forms.

2.7. Show that L-ribitol 1-phosphate is identical with D-ribitol 5-phosphate.

SOLUTION

All carbon atoms of ribitol are in the same state of oxidation, and the carbon chain can be numbered from either end. If numbered as shown in (*a*) below, and assuming C-1 was derived from an aldehyde, then the compound would be related to D-ribose and can be called D-ribitol. D-Ribitol 5-phosphate would be as shown in (*b*). However, if C-5 was derived from an aldehyde, then the compound would be related

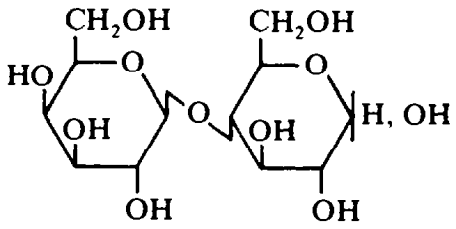
to L-ribose [turn the structure in (*a*) upside down] and called L-ribitol. L-Ribitol 1-phosphate would be as shown in (*c*), clearly the same as the phosphate ester shown in (*b*).



(*a*) (*b*) (*c*)

THE GLYCOSIDIC BOND

2.8. The disaccharide shown is lactose, the carbohydrate of mammalian milk. Give (*a*) its full name and (*b*) its abbreviated name.



Lactose

SOLUTION

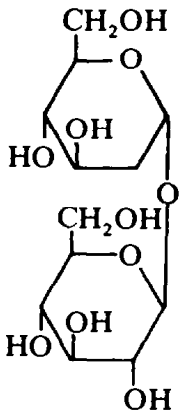
(*a*) β-Galactopyranosyl-(1 → 4)-glucopyranose

(*b*) β-Gal-(1 → 4)-Glc

2.9. Give the structure for a disaccharide, containing only glucose, that is nonreducing.

SOLUTION

For the disaccharide to be nonreducing, the anomeric carbons of both glucose molecules must be linked via a glycosidic bond. The product is called *trehalose*, and it exists in three isomeric forms in which the glucose molecules are linked α,α; β,β; or α,β (shown). α,β-Trehalose is α-D-glucopyranosyl-(1 → 1)-β-D-glucopyranose.



α,β-Trehalose

POLYSACCHARIDES

2.10. What products would result from the partial acid hydrolysis of chitin?

SOLUTION

The $\beta(1\rightarrow 4)$ linkage is relatively resistant to hydrolysis, so the major site of hydrolysis is that of the *N*-acetyl group to yield poly(α-D-glucosamine) and acetic acid. This material is also known as "chitosan".

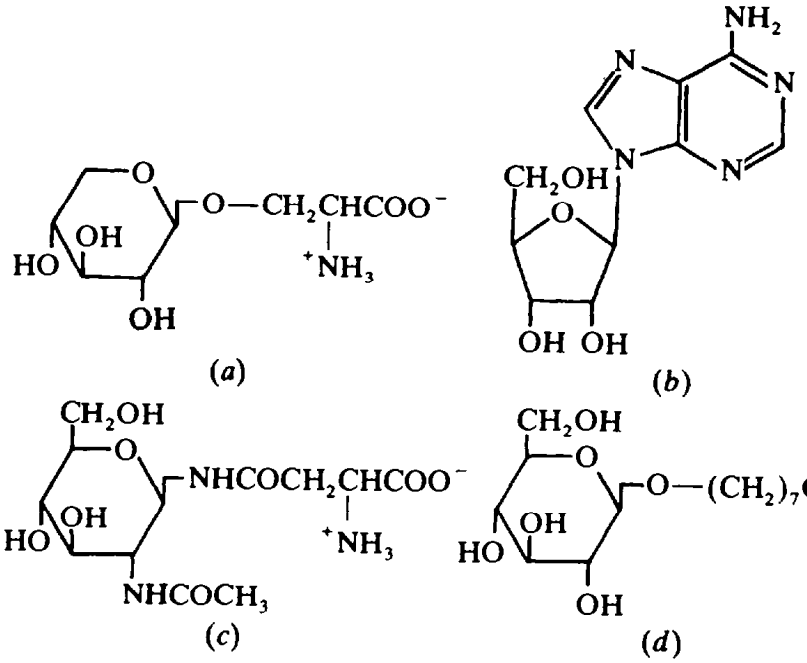
2.11. How could one determine the degree of polymerization of a sample of amylose (i.e., the average number of glucose residues per chain)?

SOLUTION

Amylose has only a single reducing group per polysaccharide chain: the group at carbon 1 of the last residue in the polymer. Accurate and sensitive measurement of the concentration of reducing groups in the sample would provide an estimate of the number of chains. Analysis of the total glucose concentration would allow an estimate of the degree of polymerization from the ratio of the concentration of total glucose to that of reducing groups.

Supplementary Problems

2.12. Identify by class and by name the monosaccharides that occur in the following glycosides. Give the configuration of each glycosidic bond.



2.13. Which of the following are reducing sugars?

(*a*)	Galactose	(*e*)	Xylose	(*i*)	Glucosamine
(*b*)	β-Methylgalactoside	(*f*)	Fructose	(*j*)	Gluconic acid
(*c*)	Maltose	(*g*)	Rhamnose	(*k*)	Glucitol
(*d*)	Mannose	(*h*)	Ribose		

2.14. (*a*) Draw the structure of any β-D-aldoheptose in the pyranose form. (*b*) Draw the structures of the anomer, the enantiomer, an epimer.

2.15. A sugar ($C_5H_{10}O_5$) was treated by a method that reduces aldehyde groups and gave a product that was optically inactive. Assuming the sugar was D, what are two possible structures of the product?

2.16. An aqueous solution of D-galactose has an $[\alpha]_D^{25}$ of $+80.2°$ after standing for some hours. The specific rotations of pure α-D-galactose and β-D-galactose are $+150.7°$ and $+52.8°$, respectively. Calculate the proportions of α- and β-D-galactose in the equilibrium mixture.

2.17. The specific rotation of maltose in water is $+138°$. What would be the concentration of a maltose solution that had an optical rotation of $+23°$ if a polarimeter tube 10 cm long was used?

2.18. An enzyme known as *invertase* catalyzes the hydrolysis of sucrose to an equimolar mixture of D-glucose and D-fructose. During the hydrolysis, the optical rotation of the solution changes from (+) to (−). What can you conclude from this observation?

2.19. Nojirimycin (5-amino-5-deoxy-D-glucose) is an antibiotic used in studies of the biosynthesis of glycoproteins. Write its open-chain and pyranose ring structures.

2.20. There are two possible chair conformers of β-D-glucopyranose. How many boat conformers are possible?

2.21. Show that in β-L-glucopyranose all the substituents on the ring carbon atoms are equatorial.

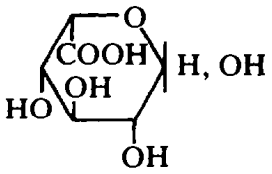
2.22. Write the preferred conformation for α,α-trehalose (see Prob. 2.9), the carbohydrate found in the hemolymph of insects.

2.23. Erythritol is the reduction product of D-erythrose. Why is the prefix D-omitted from its name?

2.24. Galactitol, the reduced form of D-galactose, is the toxic by-product that accumulates in persons suffering from galactosemia. Write its structure.

2.25. Write the open-chain structures for (*a*) D-gluconic acid and (*b*) D-glucuronic acid.

2.26. L-Iduronic acid is part of the structure of some polysaccharides found in connective tissue. How could this sugar be formed in a one-step reaction from D-glucuronic acid?



2.27. Write the Fischer projection formulas for the sugar derivatives named in Example 2.20.

2.28. The name *methylpentose* is sometimes used to describe the sugars L-fucose and L-rhamnose. Is this name valid? Explain your answer.

2.29. Are α-methyl-D-glucopyranoside and β-methyl-D-glucopyranoside anomers, isomers, or conformers?

Chapter 3

Amino Acids and Peptides

3.1 AMINO ACIDS

Amino Acids Found in Proteins

All proteins are composed of *amino acids* linked into a linear sequence by *peptide bonds* between the amino group of one amino acid and the *carboxyl* group of the preceding amino acid. The amino acids found in proteins are all *α-amino acids*; i.e., the amino and carboxyl groups are both attached to the same carbon atom (the α-carbon atom; Fig. 3-1). The α-carbon atom is a potential chiral center, and except when the —R group (or *side chain*) is H, amino acids display optical activity. All amino acids found in proteins are of the L configuration, as indicated in Fig. 3-1.

There are 20 different amino acids used in the synthesis of proteins; these amino acids are listed in Table 3.1, which also contains the two commonly used symbols for each amino acid. The three-letter symbols are easier to remember, but the single-letter symbols are often used in writing long sequences. In many proteins some of the amino acids are modified after incorporation into proteins; e.g., in collagen, a hydroxyl group is added to each of several proline residues to yield *hydroxyproline* residues. With the exception of proline, the α-amino acids that are incorporated into proteins can be represented by the formula shown in Fig. 3-1.

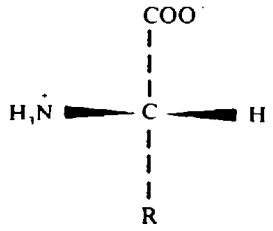


Fig. 3-1 General structure of α-amino acids in the L configuration; R is one of over 20 different chemical groups.

The side chains of the amino acids do not form a natural series, and thus, there is no easy way to learn their structures. It is useful to classify them according to whether they are polar or nonpolar, aromatic or aliphatic, or *acidic* or *basic*, although these classifications are not mutually exclusive. Tyrosine, for example, can be considered to be both aromatic and polar, although the polarity introduced by a single hydroxyl group in this aromatic compound is somewhat feeble.

The aromatic amino acids absorb light strongly in the ultraviolet region. Use may be made of this in determining the concentration of these amino acids in solution. The *Beer-Lambert* relationship states that the absorbance of light at a given wavelength by a substance in solution is proportional to its concentration C (in mol L^{-1}) and the length l (in cm) of the light path in the solution:

$$A = \varepsilon Cl \qquad (3.1)$$

where A is the absorbance of the solution and ε is the molar absorbance coefficient. The absorbance

Table 3.1 Amino Acids Used in Protein Synthesis, Grouped according to Chemical Type

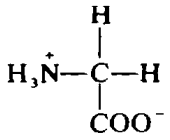
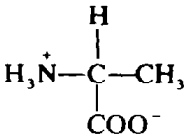
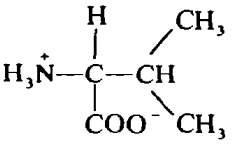
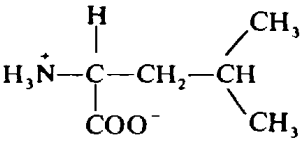
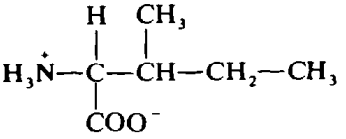
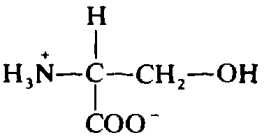
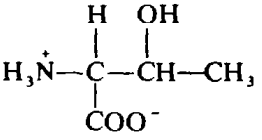
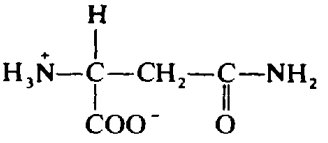
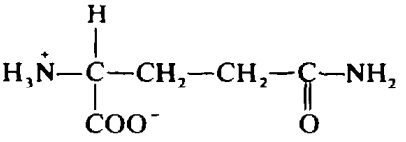
Structure	Name	Abbreviation	
1. Neutral Amino Acids (*a*) *Nonpolar, aliphatic*			
	Glycine	Gly	G
	Alanine	Ala	A
	Valine	Val	V
	Leucine	Leu	L
	Isoleucine	Ile	I
(*b*) *Polar, aliphatic*			
	Serine	Ser	S
	Threonine	Thr	T
	Asparagine	Asn	N
	Glutamine	Gln	Q

Table 3.1 (*Cont.*)

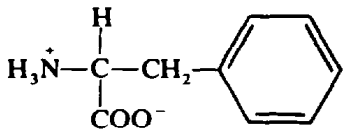
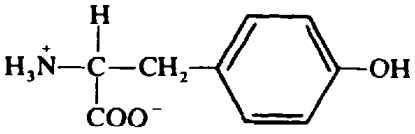
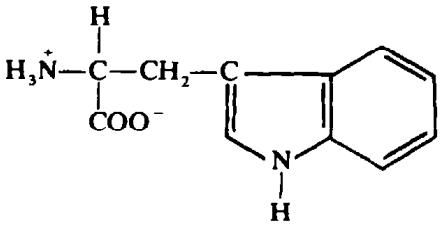
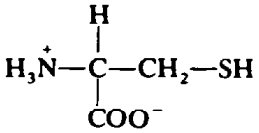
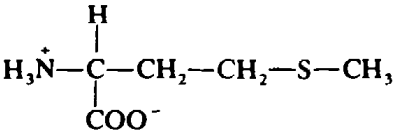
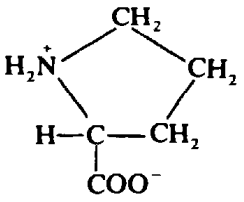
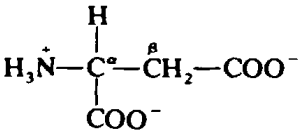
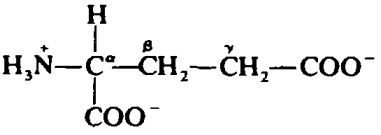
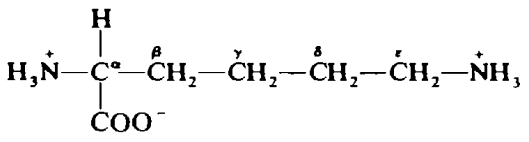
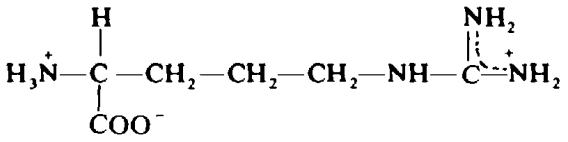
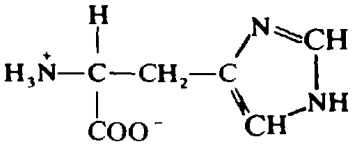
Structure	Name	Abbreviation	
(*c*) *Aromatic*			
	Phenylalanine	Phe	F
	Tyrosine	Tyr	Y
	Tryptophan	Trp	W
(*d*) *Sulfur-containing*			
	Cysteine	Cys	C
	Methionine	Met	M
(*e*) *Containing secondary amino group*			
	Proline	Pro	P
2. Acidic Amino Acids			
	Aspartate	Asp	D
	Glutamate	Glu	E

Table 3.1 (*Cont.*)

Structure	Name	Abbreviation	
3. Basic Amino Acids			
	Lysine	Lys	K
	Arginine	Arg	R
	Histidine	His	H

The Greek symbols indicate the nomenclature of the carbon chains in certain amino acids. The carbon atom carrying (i.e., next to) the carboxyl group is labeled α.

is defined as the logarithm of the ratio of the intensity of incident light (I_0) to that of transmitted light (I):

$$A = \log_{10} \frac{I_0}{I} \tag{3.2}$$

EXAMPLE 3.1

Given that the *molar absorbance coefficient* of tyrosine in water is 1,420 L mol^{-1} cm^{-1} at 275 nm, what is the concentration of tyrosine in a solution of path length 1 cm for which the absorbance is 0.71?

By use of the Beer-Lambert relationship:

$$C = \frac{A}{\varepsilon \cdot l} = \frac{0.71}{1{,}420 \times 1.0} = 5 \times 10^{-4} \text{ mol L}^{-1}$$

Nonprotein Amino Acids

A large number of amino acids involved in metabolism are not found in proteins; e.g., β-alanine, $^-OOC-CH_2-CH_2-NH_3^+$, is an intermediate in the synthesis of the B vitamin *pantothenic acid*, but it is not found in proteins. Although most naturally occurring amino acids are of the L configuration, some D-amino acids are found in certain antibiotics and in the cell walls of some bacteria.

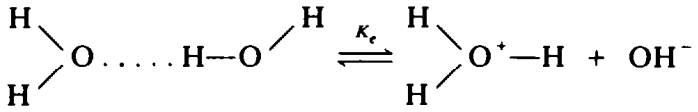
3.2 ACID-BASE BEHAVIOR OF AMINO ACIDS

Amino acids are *amphoteric* compounds; i.e., they contain both acidic and basic groups. Because of this, they are capable of bearing a net electrical charge, which depends on the nature of the solution.

The charge carried by a molecule influences the way in which it interacts with other molecules; use can be made of this property in the isolation and purification of amino acids and proteins. Therefore, it is important to have a clear understanding of the factors that influence the charge carried on amino acids.

The Ionization of Water

The major biological solvent is water, and the acid-base behavior of dissolved molecules is intimately linked with the dissociation of water. Water is a weak electrolyte capable of dissociating to a proton and a hydroxyl ion. In this process, the proton binds to an adjacent water molecule to which it is *hydrogen-bonded* (Chap. 4) to form a *hydronium ion* (H_3O^+):



In pure water at 25°C, at any instant there are 1.0×10^{-7} mol L^{-1} of H_3O^+ and an equivalent concentration of OH^- ions. It must be stressed that the proton is hardly ever "bare" in water because it has such a high affinity for water molecules. Hydrated complexes other than H_3O^+ have been suggested, but since water is so extensively hydrogen-bonded, it is difficult to identify these species experimentally, and as a simplification the hydrated proton is often written as H^+.

The dissociation of water is a rapid equilibrium process for which we can write an equilibrium constant:

$$K_e = \frac{a_{H_3O^+} \times a_{OH^-}}{(a_{H_2O})^2} \tag{3.3}$$

Since, for dilute solutions, the *activity a* (Chap. 10) of water is considered to be constant and very close to 1.0, and the activities of the solutes may be represented by their concentrations, we can define a practical constant, K_w, called the *ionic product of water*:

$$K_w = [H_3O^+] \cdot [OH^-] \tag{3.4}$$

or, as is often seen, $K_w = [H^+][OH^-]$, with the hydration of the proton ignored for simplicity. Note that the square brackets denote concentrations of species in mol L^{-1}.

At 25°C in pure water, $K_w = 10^{-14}$. Since, in pure water, $[H^+] = [OH^-]$

$$[H^+] = \sqrt{10^{-14}} = 10^{-7} \text{ mol } L^{-1}$$

In acid solution, $[H^+]$ is higher, and $[OH^-]$ is correspondingly lower, because the ionic product is constant. The value of K_w, is temperature-dependent; at 37°C, for example, $K_w = 2.4 \times 10^{-14}$.

EXAMPLE 3.2

Calculate $[OH^-]$ in aqueous solution at 25°C when $[H^+] = 0.1$ mol L^{-1}.
Since $[H^+] \cdot [OH^-] = 10^{-14}$

$$[OH^-] = \frac{10^{-14}}{10^{-1}} = 10^{-13} \text{ mol } L^{-1}$$

Acidity and pH

The Danish chemist S. P. L. Sørensen defined pH (*potentia Hydrogenii*) as:

$$pH = -\log_{10}[H^+] \tag{3.5}$$

Neutral solutions are defined as those in which $[H^+] = [OH^-]$, and for pure water at 25°C

$$pH = -\log_{10}(10^{-7}) = 7.0 \tag{3.6}$$

Note that *distilled* water normally used in the laboratory is not absolutely pure. Traces of CO_2 dissolved in it produce carbonic acid that increases the hydrogen-ion concentration to about 10^{-5} mol L^{-1}, thus rendering the pH around 5.

EXAMPLE 3.3

Calculate the pH of a 4×10^{-4} mol L^{-1} solution of HCl.

At this low concentration we may consider HCl to be completely dissociated to H^+ and Cl^-. Therefore

$$[H^+] = 4 \times 10^{-4} \text{ mol } L^{-1}$$

and

$$pH = -\log(4 \times 10^{-4}) = 3.40$$

Note that acid solutions (high H^+ concentrations) have *low* pH and alkaline solutions (high OH^- concentration and *low* H^+ concentration) have *high* pH. A 10-fold increase in $[H^+]$ corresponds to a decrease of 1.0 in pH.

Weak Acids and Bases

Acids

An *acid* is a compound capable of donating a proton to another compound (this is the so-called *Brønsted definition*). The substance CH_3COOH is an acid, *acetic acid*. However, because the dissociation of all the carboxyl groups is not complete when acetic acid is dissolved in water, acetic acid is referred to as a *weak* acid. The dissociation reaction for any weak acid of type HA in water is

$$HA + H_2O \leftrightharpoons A^- + H_3O^+$$

The dissociation reaction for acetic acid is therefore

$$\underset{\substack{H^+ \text{ donor} \\ \text{(acid)}}}{CH_3COOH} + \underset{\substack{H^+ \text{ acceptor} \\ \text{(base)}}}{H_2O} \rightleftharpoons \underset{\substack{\text{Conjugate} \\ \text{base}}}{CH_3COO^-} + \underset{\substack{\text{Conjugate} \\ \text{acid}}}{H_3O^+}$$

The donating and accepting of the proton is a two-way process. The H_3O^+ ion that is formed is capable of donating a proton back to the acetate ion to form acetic acid. This means that the H_3O^+ ion is considered to be an acid and the acetate ion is considered a base. Acetate is called the *conjugate base* of acetic acid.

The two processes of association and dissociation come to equilibrium, and the resulting solution will have a higher concentration of H_3O^+ than is found in pure water; i.e., it will have a pH below 7.0.

A measure of the *strength* of an acid is the *acid dissociation constant*, K_a:

$$K_a = \frac{a_{H_3O^+} \cdot a_{A^-}}{a_{HA} \cdot a_{H_2O}} \tag{3.7}$$

For acetic acid, K_a is therefore

$$K_a = \frac{a_{H_3O^+} \cdot a_{CH_3COO^-}}{a_{CH_3COOH} \cdot a_{H_2O}}$$

where a denotes the thermodynamic activity (Chap. 10) of the chemical species.

In dilute solutions, the concentration of water is very close to that of pure water, and the activity of pure water, by convention, is taken to be 1.0. Furthermore, in dilute solutions, the activity of solutes may be approximated by their concentrations; so we may write an expression for a *practical* acid dissociation constant:

$$K_a = \frac{[H^+][A^-]}{[HA]} \tag{3.8}$$

For acetic acid, this becomes

$$K_a = \frac{[H^+][CH_3COO^-]}{[CH_3COOH]}$$

The *larger* the value of K_a, the greater the tendency of the acid to dissociate a proton, and so the *stronger* the acid.

In a manner similar to the definition of pH, we can define

$$pK_a = -\log K_a \tag{3.9}$$

Thus, the *lower* the value of the pK_a of a chemical compound, the higher the value of K_a, and the *stronger* it is as an acid.

EXAMPLE 3.4

Which of the following acids is the stronger: boric acid, which has a $pK_a = 9.0$, or acetic acid, with a $pK_a = 4.6$?

For boric acid, $K_a = 10^{-9}$, while for acetic acid, $K_a = 10^{-4.6} = 2.5 \times 10^{-5}$. Thus, acetic acid has the greater K_a and is therefore the stronger acid.

Bases

A *base* is a compound capable of accepting a proton from an acid. When methylamine, $CH_3{-}NH_2$, dissolves in water, it accepts a proton from the water, thus leading to an increase in the OH^- concentration and a high pH.

$$\underset{\text{Base}}{CH_3{-}NH_2} + \underset{\text{Acid}}{H_2O} \rightleftharpoons \underset{\text{Conjugate acid}}{CH_3{-}NH_3^+} + \underset{\text{Conjugate base}}{OH^-}$$

As in the case of acetic acid, as the concentration of OH^- increases, the reverse reaction becomes more significant and the process eventually reaches equilibrium.

We can write an expression for a *basicity constant*, K_b:

$$K_b = \frac{[CH_3{-}NH_3^+][OH^-]}{[CH_3{-}NH_2]} \tag{3.10}$$

However, the use of this constant can be confusing, as we would need to keep track of two different types of constant, K_a and K_b. Since chemical equilibrium is a two-way process, it is perfectly correct, and more convenient, to consider the behavior of bases from the point of view of their *conjugate acid*. The latter can be considered to donate a proton to water:

$$CH_3{-}NH_3^+ + H_2O \rightleftharpoons CH_3{-}NH_2 + H_3O^+$$

and

$$K_a = \frac{[CH_3{-}NH_2][H_3O^+]}{[CH_3{-}NH_3^+]} \tag{3.11}$$

Of course, K_a and K_b are related as follows:

$$K_a \cdot K_b = \frac{[CH_3{-}NH_2][H_3O^+]}{[CH_3{-}NH_3^+]} \cdot \frac{[CH_3{-}NH_3^+][OH^-]}{[CH_3{-}NH_2]} = [H_3O^+][OH^-]$$

i.e.,

$$K_a \cdot K_b = K_w \tag{3.12}$$

In other words, if we know K_a for the conjugate acid, we can calculate K_b for the base. A base is thus characterized by a *low* value of K_a for its conjugate acid.

Buffers

A mixture of an acid and its conjugate base is capable of resisting changes in pH when small amounts of additional acid or base are added. Such a mixture is known as a *buffer*.

Consider again the dissociation of acetic acid:

$$CH_3COOH + H_2O \rightleftharpoons CH_3COO^- + H_3O^+$$

Additional acid causes recombination of H_3O^+ and CH_3COO^- to form acetic acid, so that a buildup of H_3O^+ is resisted. Conversely, addition of NaOH causes dissociation of acetic acid to acetate, reducing the fall in H_3O^+ concentration.

This behavior can be quantified by taking logarithms of both sides of Eq. (*3.8*):

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

$$\log K_a = \log [H^+] + \log [A^-] - \log [HA]$$

Multiplying both sides by -1, we get

$$-\log K_a = -\log [H^+] - \log [A^-] + \log [HA]$$

From Eq. (*3.9*), $-\log K_a = pK_a$, and from Eq. (*3.5*), $-\log [H^+] = pH$. By substitution, we therefore get

$$pK_a = pH + \log \frac{[HA]}{[A^-]} = pH + \log \frac{[acid]}{[conjugate\ base]} \qquad (3.13a)$$

or

$$pH = pK_a + \log \frac{[base]}{[conjugate\ acid]} \qquad (3.13b)$$

These are two forms of the *Henderson-Hasselbalch equation*. This useful relationship enables us to calculate the composition of buffers that have a specified pH. Note that if $[HA] = [A^-]$, then $pH = pK_a$.

EXAMPLE 3.5

(*a*) Calculate the pH of a solution containing 0.1 mol L^{-1} acetic acid and 0.1 mol L^{-1} sodium acetate. The pK_a of acetic acid is 4.7. (*b*) What would be the pH value after adding 0.05 mol L^{-1} NaOH? (*c*) Compare the latter value with the pH of a simple solution of 0.05 mol L^{-1} NaOH.

(*a*) Using the Henderson-Hasselbalch equation (*3.13b*),

$$pH = pK_a + \log \frac{[base]}{[acid]}$$

Therefore

$$pH = 4.7 + \log \frac{0.1}{0.1} = 4.7 + 0 = 4.7$$

(*b*) On adding 0.05 *M* NaOH, the concentration of undissociated acetic acid falls to 0.05 *M*, while the acetate concentration rises to 0.15 *M*. Therefore,

$$pH = 4.7 + \log \frac{0.15}{0.05} = 4.7 + \log 3 = 4.7 + 0.48 = 5.18$$

(*c*) The pH of 0.05 mol L^{-1} NaOH is $-\log [H^+]$ in the solution. If we assume the NaOH is fully dissociated in water, the value of $[OH^-]$ is 0.05 mol L^{-1}. The known value of the ionic product of water is 10^{-14} (mol $L^{-1})^2$; therefore,

$$[H^+] = 10^{-14}/0.05$$

and

$$pH = 12.7$$

In comparison with this, even after adding the same amount of alkali, the buffered solution changes pH only slightly.

Acid-Base Behavior of Simple Amino Acids

Many biological molecules have more than one dissociable group. The dissociation of one group can have profound effects on the tendency for dissociation of the other groups. Amino acids, containing both carboxyl and amino groups, illustrate this phenomenon. In water, the carboxyl group tends to dissociate a proton, while the amino group binds a proton. Both reactions can therefore proceed largely to completion, with no buildup either of H_3O^+ or OH^-. An important result is that amino acids carry both a negative and a positive charge in solution near neutral pH; in this state the compound is said to be a *zwitterion*.

Titration of Amino Acids

A way of examining the zwitterionic behavior of amino acids is to study their *titration*. Suppose, for example, that we begin with a solution of glycine hydrochloride, in which both groups are in their acidic forms. Addition of sodium hydroxide brings about an increase in the pH of the solution, and at the same time, dissociated protons react with the added hydroxyl ions to form water thus allowing further dissociation to take place, as shown in Fig. 3-2.

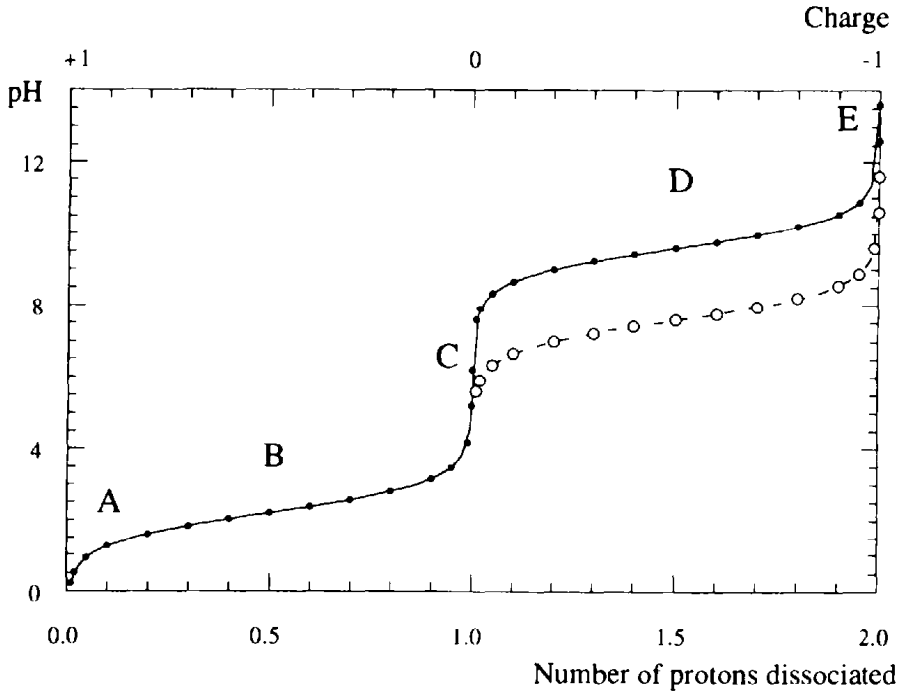


Fig. 3-2 Titration of 10 mmol glycine·HCl with NaOH solution. The broken line shows the titration in the presence of formaldehyde.

The curve in Fig. 3-2 shows two distinct branches, one for each acidic group on glycine hydrochloride. At point A, both groups are in the acid (protonated) form: $HOOC{-}CH_2{-}NH_3^+$ and the charge carried by the molecule is +1. After addition of 10 mmol of NaOH (point C, near pH 6.0), one group is almost completely deprotonated, and the molecule carries no net charge. After an additional 10 mmol of NaOH, the second group is also deprotonated (point E), and the form of the molecule will be $^-OOC{-}CH_2{-}NH_2$, with a charge of −1.

In the vicinity of points B and D, the pH changes least for a given amount of added NaOH; i.e., the solution is acting as a buffer. At point B, pH 2.3, half of the first group has been titrated to its conjugate base and half remains in its acid form; i.e., [acid] = [conjugate base]. At this point,

then, from Eq. (*3.13*), $pK_a = pH$, or for this group, $pK_a = 2.3$. The first group to titrate must be a relatively strong acid, and this group would be expected to be the carboxyl group. Similar considerations lead to a value of 9.6 for the pK_a of the second group, the amino group.

EXAMPLE 3.6

What fraction of the first group (the carboxyl) has been titrated by pH 6.3?
Using the Henderson-Hasselbalch equation (*3.13b*):

$$pH = pK_a + \log \frac{[\text{base}]}{[\text{acid}]}$$

At pH 6.3, for a group with a pK_a of 2.3:

$$\log \frac{[\text{base}]}{[\text{acid}]} = 6.3 - 2.3 = 4.0$$

Therefore, [base] = $10^4 \times$ [acid]. This result indicates that only 1 in 10,000 of the carboxyl groups (0.01 percent) is protonated at pH 6.3.

The pH at which the molecule carries no net charge is called the *isoelectric point*. For glycine the isoelectric point is pH 6.0. Of course, in a solution of glycine at pH 6, at any instant there will be some molecules that exist as $COOH{-}CH_2{-}NH_3^+$, and an equal number as $COO^-{-}CH_2{-}NH_2$, and even fewer of $COOH{-}CH_2{-}NH_2$. Because of the symmetry of the titration curve around the isoelectric point, it is possible to calculate the pH of the isoelectric point, given the individual pK_a values.

At the isoelectric point:

$$pH_I = \frac{pK_{a1} + pK_{a2}}{2} \tag{3.14}$$

The Formol Titration

Chemical modification can be used to verify the assignment of pK_a values in titrations of amino acids. For example, by titrating in the presence of formaldehyde, we can show that pK_{a_1} in Fig. 3-2 belongs to the carboxyl group, and pK_{a2} to the amino group.

EXAMPLE 3.7

Formaldehyde (HCHO) reacts reversibly with unchanged amino groups as follows:

$$R{-}NH_2 + 2HCHO \rightleftharpoons R{-}N\begin{matrix}\diagup CH_2{-}OH \\ \diagdown CH_2{-}OH\end{matrix}$$

Thus, the pK_a of amino groups is altered in formaldehyde solutions, while the carboxyl groups remain unaffected. It is found that the pK_{a1} of amino acids is unchanged in formaldehyde and therefore represents the carboxyl group, while pK_{a2} is lowered (see the broken line in Fig. 3-2). The extent of lowering depends on the concentration of formaldehyde. Thus, pK_{a2} reflects the dissociation of the amino group.

In general, among simple compounds, the carboxyl group is a stronger acid than the $-NH_3^+$ group. Carboxyl groups tend to have pK_a values below 5 and $-NH_3^+$ groups above 7.

Question: Why is the pK_a of the glycine carboxyl (2.3) less than the pK_a of acetic acid (4.7)?

In glycine solutions at pH values below 6, the amino group is present in the positively charged form. This positive charge stabilizes the negatively charged carboxylate ion by electrostatic interaction.

This means that the carboxyl group of glycine will *lose* its proton more readily and is therefore a stronger acid (with a lower pK_a value).

Acidic and Basic Amino Acids

Some of the amino acids carry a *prototropic* side chain: e.g., aspartic acid and glutamic acid have an extra carboxyl, histidine has an imidazole, lysine has an amino, and arginine has a guanidino group. The structures of these side chains are shown in Table 3.1, and their pK_a values are listed in Table 3.2.

The titration curves of these amino acids have an extra inflection, as shown for glutamic acid in Fig. 3-3.

Table 3.2. pK_a Values of Some Amino Acids

Amino Acid	pK_{a1} (α-COOH)	pK_a (α-NH_3^+)	pK_{aR} (side chain)
Glycine	2.3	9.6	—
Serine	2.2	9.2	—
Alanine	2.3	9.7	—
Valine	2.3	9.6	—
Leucine	2.4	9.6	—
Aspartic acid	2.1	9.8	3.9
Glutamic acid	2.2	9.7	4.3
Histidine	1.8	9.2	6.0
Cysteine	1.7	10.8	8.3
Tyrosine	2.2	9.1	10.1
Lysine	2.2	9.0	10.5
Arginine	2.2	9.0	12.5

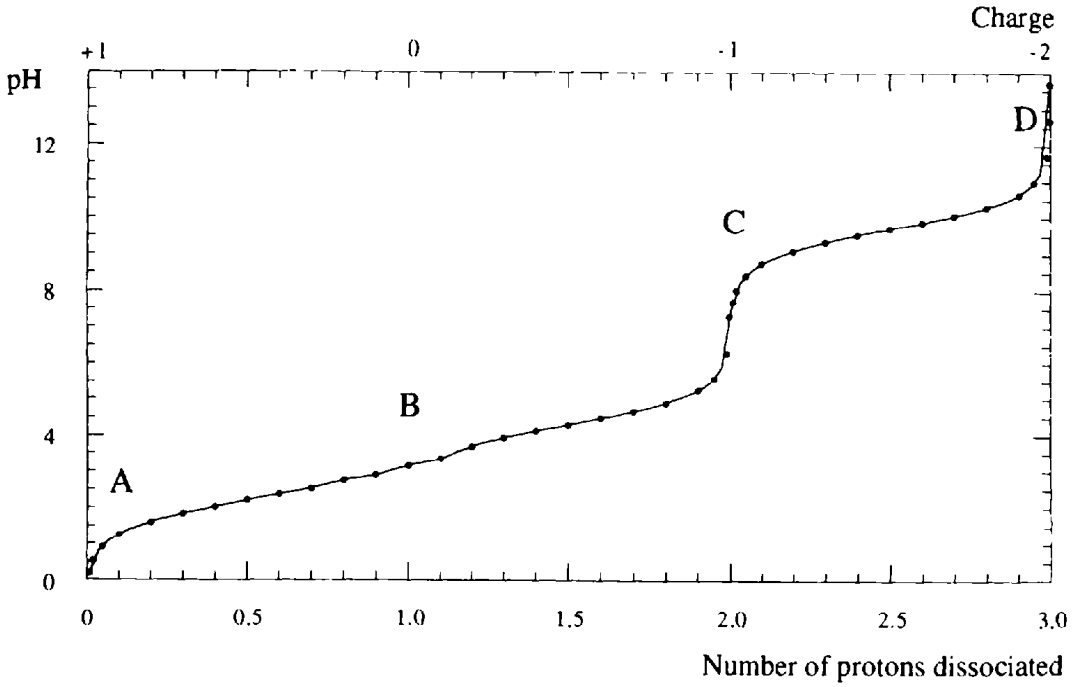


Fig. 3-3 Titration of 10 mmol glutamic acid hydrochloride by NaOH.

Charge on Polybasic Amino Acids

Recall that for amino acids, the various groups have a charge that depends on the pH of the solution:

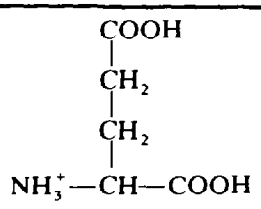
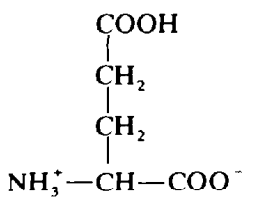
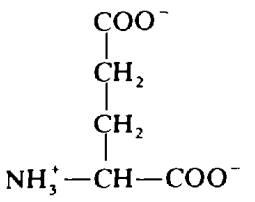
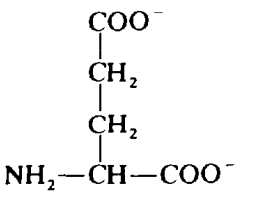
1. The protonated forms of the carboxyl groups and the tyrosine side chain are *uncharged*, while the deprotonated forms are negatively charged, or *anionic*.
2. The protonated forms of the amino group, the *imidazolium* side chain of histidine and the *guanidinium* group of arginine, are positively charged (or *cationic*), while the deprotonated forms are uncharged.

At pH values *below* the pK_a of a group, the solution is more acidic, and the protonated form predominates. As the pH is raised above the pK_a of a group, that group loses its proton; i.e., the $-COOH$ group becomes the negative $-COO^-$, while the positively charged acid form of the amino group, $-NH_3^+$, becomes *uncharged* NH_2.

Isoelectric Point

By using the above arguments, we can consider the example of the isoelectric point of glutamic acid, with the aid of the pK_a values given in Table 3.2.

At pH values less than 2.2, *all* the groups are protonated. The predominant forms of glutamic acid, present at different points during the titration shown in Fig. 3-3, are given below:

Point	Charged Form	Charge
A (pH 1)		+1
B (pH 3.25)		0
C (pH 7)		−1
D (pH 12)		−2

As the pH is increased through 2.2, the α-carboxyl dissociates to yield the zwitterion near pH 3. As the pH is further increased through 4.3, the side-chain carboxyl dissociates to yield a molecule with

two negative and one positive charges (i.e., a net charge of −1). Note that near pH 7 (i.e., in neutral solution), both glutamic acid and aspartic acid exist as glutamate and aspartate, respectively, with ionized side chains. Finally, as the pH rises through the pK_{a2}, the amino group dissociates to yield the species with a charge of −2. Thus, the isoelectric point is flanked by the two pK_a values of 2.2 and 4.3.

So, it can be shown that for glutamic acid,

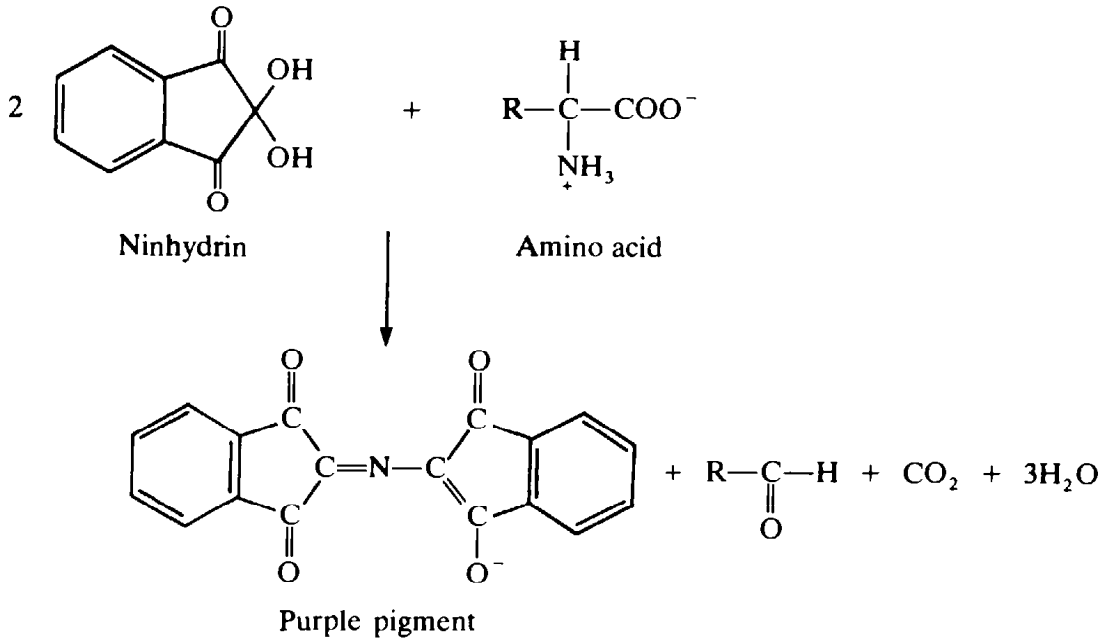
$$pH_I = (pK_{a1} + pK_{aR})/2 = (2.2 + 4.3)/2 = 6.5/2 = 3.25.$$

Thus, as a general rule, the isoelectric point of an amino acid is the average of the pK_a values of the protonation transitions on *either side* of the isoelectric species; this implies a knowledge of the charges of the various forms.

3.3 AMINO ACID ANALYSIS

After hydrolysis of proteins to amino acids (usually in concentrated HCl), the amino acids can be separated from each other by means of *ion-exchange chromatography*. Three buffers of successively higher pH are used to *elute* the amino acids from the chromatography column. The order of elution depends on the charge carried by the amino acid. The basic amino acids (lysine, histidine, and arginine) bind most tightly to the negatively charged ion-exchange resin. By using this technique, it is possible to determine which amino acids occur in a given protein. Their *relative abundance* can also be determined by measuring the concentration of each amino acid. The compound *ninhydrin* reacts with amino acids to form a purple derivative. By measuring the absorbance at 570 nm of the purple solution so formed, the relative concentrations of each amino acid can be determined (Fig. 3-4).

Reaction of amino acids with ninhydrin:



Although proline does not give a purple color when treated with ninhydrin, it does yield a weaker yellow pigment, which can be quantified. Proline can be located readily in a chromatogram during amino acid analysis because the relative heights of the 570- and 440-nm absorbance profile are reversed compared with the other amino acids.

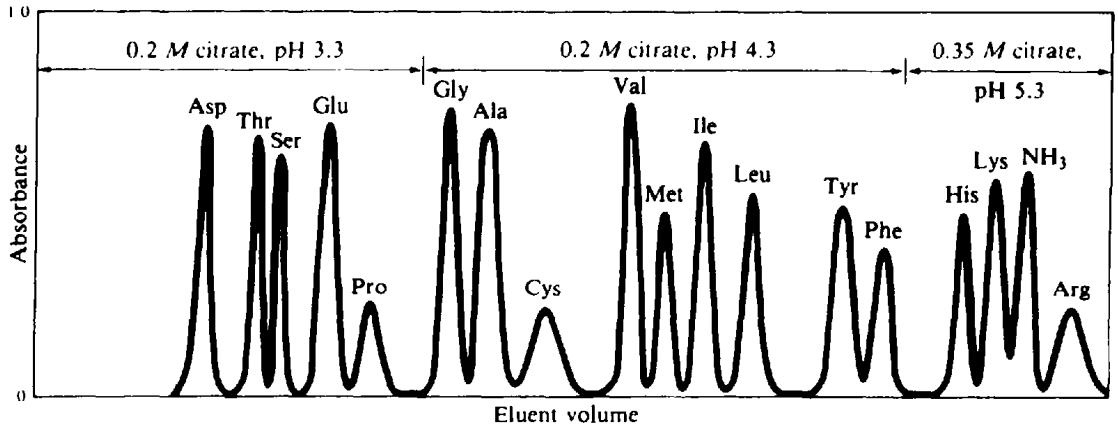


Fig. 3-4 Separation of amino acids by ion-exchange chromatography. The areas under the peaks are proportional to the amounts of amino acids in the solution.

EXAMPLE 3.8

A peptide was hydrolyzed, and the resulting solution was examined by amino acid analysis. The following data were obtained:

Amino Acid	μmol
Asp	1.21
Ser	0.60
Gly	1.78
Leu	0.58
Lys	0.61

Determine the empirical formula of the peptide.

We may take leucine as a convenient reference compound and determine the proportion of each of the amino acids relative to leucine. This yields the following stoichiometry:

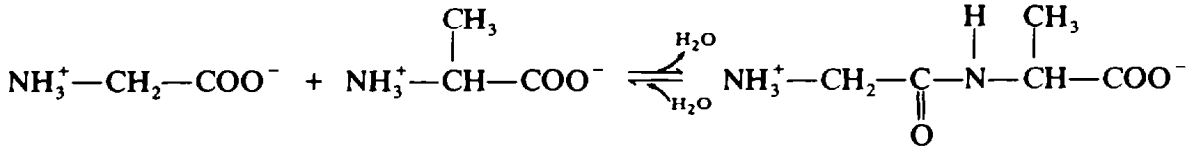
Amino Acid	Relative Concentration
Asp	2.086
Ser	1.03
Gly	3.07
Leu	1.00
Lys	1.05

Since there must be a whole number of each amino acid in the peptide, these values can be rounded off to integers to give the most likely empirical formula: Asp_2, Ser, Gly_3, Leu, Lys.

3.4 THE PEPTIDE BOND

In protein molecules, α-amino acids are linked in a linear sequence. The α-carboxyl group of one amino acid is linked to the α-amino group of the next through a special amide bond known as

a *peptide bond*. The peptide bond is formed by a *condensation* reaction, requiring the input of energy:



Glycine Alanine Glycylalanine

Note that the acidic and basic character of the carboxyl and amino groups taking part in forming the bond are lost after condensation. The hydrolysis of the peptide bond to free amino acids is a spontaneous process, but is normally very slow in neutral solution.

Nomenclature

In naming peptides, we start with the amino acid that has the free α-NH_3^+ (the *N terminal*) and replace the *-ine* endings (except the last one) with the ending *-yl*. The amino acids in the peptide are called *residues*, since they are the residues left after the removal of water during peptide bond formation.

EXAMPLE 3.9

Distinguish between the dipeptides glycylalanine and alanylglycine.

We start naming a peptide from the amino terminus; thus, glycylalanine has a free α-amino group on the glycine residue, while the free carboxyl group is that of the alanine residue.

Glycylalanine and alanylglycine are examples of *sequence isomers*; they are composed of the same amino acids, but they are combined in different sequences.

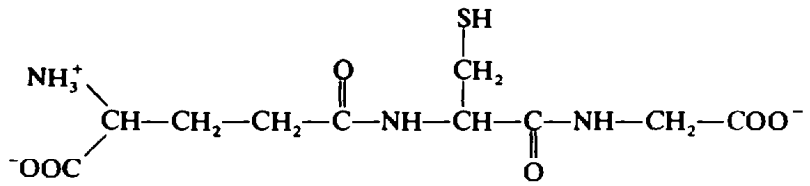
Peptides

Compounds of two amino acids linked by a peptide bond are known as *dipeptides*; those with three amino acids are called *tripeptides*, and so on. *Oligopeptides* contain an unspecified but small number of amino acid residues, while *polypeptides* comprise larger numbers. Natural polypeptides of 50 or more residues are generally referred to as *proteins* (Chap. 4).

EXAMPLE 3.10

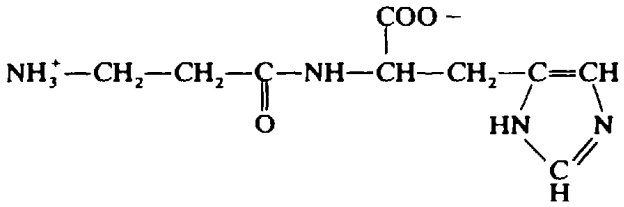
The following peptides are not derived from proteins. How could this be deduced from an inspection of their structures?

(*a*) Glutathione (γ-L-glutamyl-L-cysteinylglycine)



The peptide bond between glutamic acid and cysteine is through the γ-carboxyl of glutamic acid, not the α-carboxyl group. All natural proteins are composed with α-peptide links (unless there have been postsynthetic modifications).

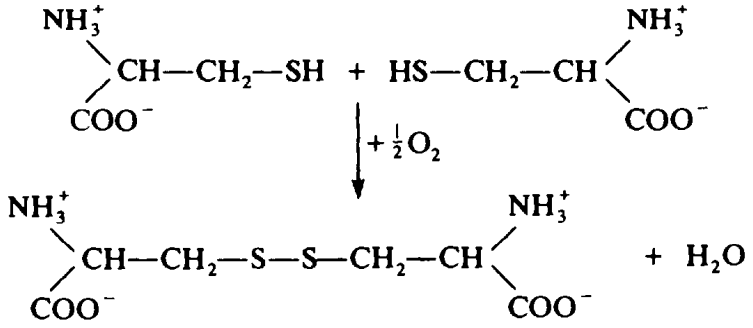
(*b*) Carnosine



This compound contains β-alanine. Naturally occurring proteins are composed of α-amino acids only.

3.5 REACTIONS OF CYSTEINE

The side chain of cysteine is important because of the possibility of its oxidation to form the disulfide-bridged amino acid *cystine*:

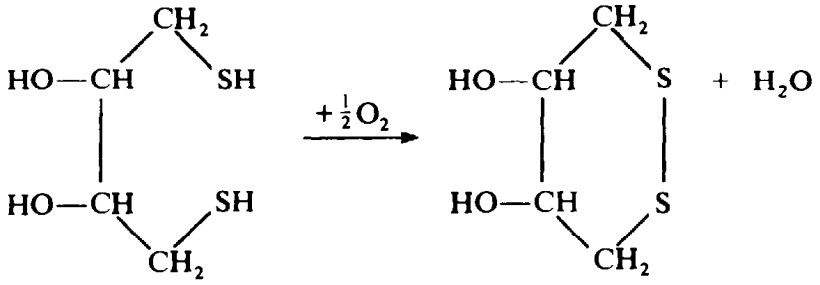


Cystine

Disulfide bridges are often found in proteins if the cysteine side chains are close enough in the *tertiary structure* (Chap. 4) to form a bridge. In addition, oxidation of free —SH groups on the surface of some proteins can cause two different molecules to be linked covalently by a disulfide bridge. This process may be biologically undesirable, and cells frequently contain reducing agents that prevent or reverse this reaction. The most common of these agents is *glutathione*; it can reduce the oxidized disulfide back to the sulfhydryl form, becoming itself oxidized in the process. Cells contain other reducing systems linked to metabolism that can then re-reduce the glutathione.

Disulfide links can be broken in the laboratory by reagents similar to glutathione that carry a free —SH group. The most common of these is *mercaptoethanol*, $HO{-}CH_2{-}CH_2{-}SH$.

There are more powerful disulfide reducing agents (lower *standard redox potential*, Chap. 10); e.g., *dithiothreitol* carries two sulfhydryl groups, and oxidation of these causes ring closure, forming a very stable disulfide. As a result, dithiothreitol is several orders of magnitude more powerful as a reducing agent than mercaptoethanol.



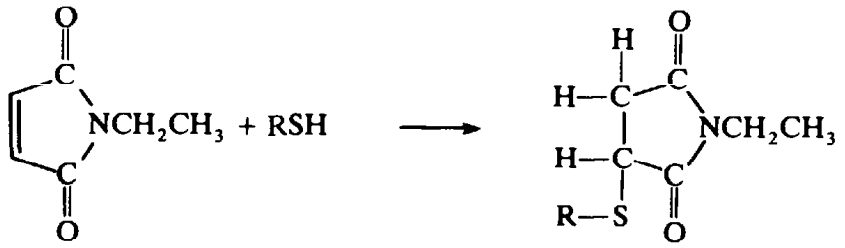
Reduced Dithiothreitol Oxidized Dithiothreitol

Often it is necessary to prevent the oxidation of sulfhydryl groups, as the presence of disulfide links could lead to insolubility of a protein or to inability to determine its sequence. It is possible to *block* the reactive —SH groups with a range of chemical reagents.

1. *Iodoacetate.* This reagent forms an *S*-carboxymethyl derivative of cysteine residues:

$$R{-}SH + I{-}CH_2{-}COO^- \rightarrow RS{-}CH_2{-}COO^- + HI$$

2. N-*Ethylmaleimide.* The reaction with *N*-ethylmaleimide results in a loss in absorbance of the reagent at 305 nm; this characteristic can be used to measure the extent of reaction:



N-Ethylmaleimide

EXAMPLE 3.11

A solution of a protein containing 2 mg in 1 mL was treated with an excess of *N*-ethylmaleimide. During the reaction in a cuvette of path length 1 cm, the absorbance at 305 nm fell from 0.26 to 0.20. Given a molar absorbance coefficient for *N*-ethylmaleimide of 620 L mol^{-1} cm^{-1}, (*a*) calculate the concentration of sulfhydryl groups in the solution; (*b*) what is the molecular weight of the protein, assuming there is only one cysteine residue per molecule?

(*a*) The concentration of sulfhydryl groups originally in the sample is equal to the decline in *N*-ethylmaleimide concentration. The Beer-Lambert law (Eq. *3.1*) can be used to determine the concentration change, given the change in absorbance:

$$A = \varepsilon C l$$

$$\Delta C = \frac{\Delta A}{\varepsilon l}$$

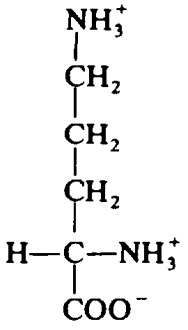
$$= \frac{0.26 - 0.20}{620 \times 1.0} = \frac{0.06}{620} = 9.8 \times 10^{-5}\ \text{mol L}^{-1}$$

(*b*) The protein concentration is 2 mg mL^{-1}, or 2 g L^{-1}. Since there is a single —SH per protein molecule, then the molar protein concentration would also be 9.8×10^{-5} mol L^{-1}. Therefore, molar weight = $2/(9.8 \times 10^{-5})$ g mol^{-1}; i.e., M_r = 20,400.

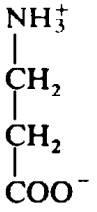
Solved Problems

AMINO ACIDS

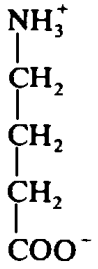
3.1. Which of the following compounds are α-amino acids?



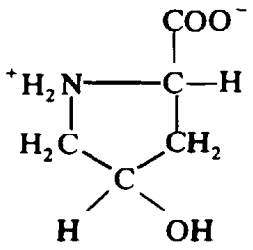
(*a*) Ornithine



(*b*) β-Alanine



(*c*) γ-Aminobutyrate



(*d*) Hydroxyproline

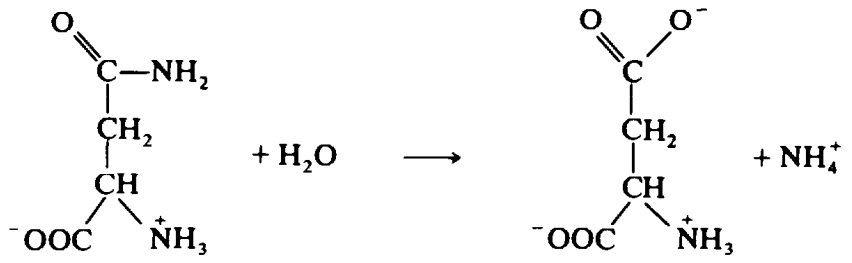
SOLUTION

Ornithine (*a*) and hydroxyproline (*d*) are both α-amino acids, because an amino group is attached to the same carbon atom as that which carries a carboxyl. Although ornithine is not used in protein synthesis, it is an intermediate in the urea cycle (Chap. 15). β-Alanine (*b*) and γ-aminobutyrate (*c*) have their amino group attached to a carbon atom different from that bearing the carboxyl, and are a β- and a γ-amino acid, respectively. Strictly speaking proline is a *secondary* amino acid, because the amino nitrogen is covalently connected to the side chain. It is sometimes referred to as an *imino acid*.

3.2. Which amino acids can be converted into different amino acids by mild hydrolysis, with the liberation of ammonia?

SOLUTION

Both glutamine and asparagine have *amide* side chains. Amides can be hydrolyzed to yield the carboxyl and free ammonia.



Asparagine **Aspartate**

3.3. Why is phenylalanine very poorly soluble in water, while serine is freely water-soluble?

SOLUTION

The aromatic side chain of phenylalanine is nonpolar, and its solvation by water is accompanied by a loss of entropy and is therefore unfavorable. On the other hand, the side chain of serine carries a polar hydroxyl group that allows hydrogen bonding with water.

3.4. Many proteins absorb ultraviolet light strongly at 280 nm, yet gelatin does not. Suggest an explanation.

SOLUTION

The ultraviolet absorbance of many proteins is due to the presence of amino acids with aromatic side chains. Gelatin, a derivative of collagen (Chap. 4), has an unusual composition, with a low proportion of aromatic amino acids.

ACID-BASE BEHAVIOR OF AMINO ACIDS

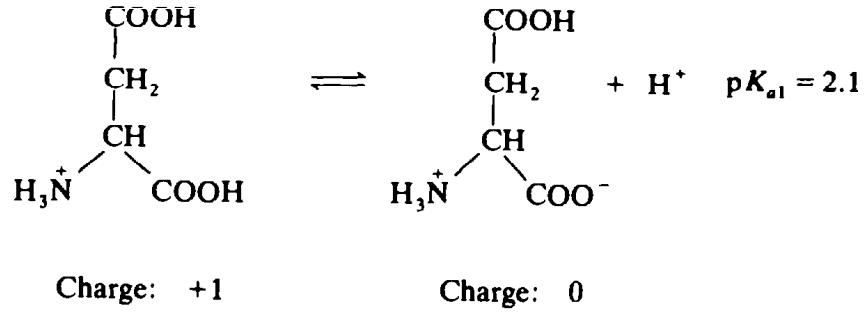
3.5. Write the conjugate bases of the following weak acids: (*a*) CH_3COOH; (*b*) NH_4^+; (*c*) $H_2PO_4^-$; (*d*) $CH_3NH_3^+$

SOLUTION

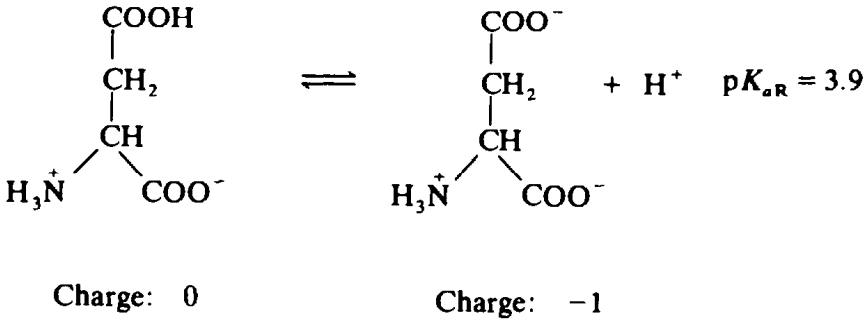
Each acid forms its conjugate base by loss of a proton. The conjugate bases are thus: (*a*) CH_3COO^-; (*b*) NH_3; (*c*) HPO_4^{2-}; (*d*) CH_3NH_2

3.6. Write the equations for the ionization equilibria of aspartic acid. Indicate the net charge carried on each species.

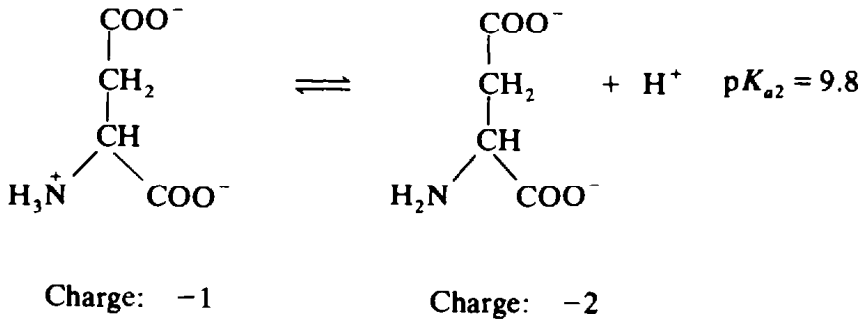
SOLUTION



Charge: +1 Charge: 0



Charge: 0 Charge: −1



Charge: −1 Charge: −2

3.7. Calculate the isoelectric point (pH_I) of aspartic acid, using the information from Prob. 3.6.

SOLUTION

Since the isoelectric species is flanked by transitions whose pK_a values are 2.1 and 3.9, respectively,

$$pH_I = \frac{2.1 + 3.9}{2} = 3.0$$

3.8. Calculate the pH of 0.1 *M* acetic acid ($pK_a = 4.7$).

SOLUTION

Acetic acid, although a weak acid, is a stronger acid than water. Therefore, as a first approximation, we can ignore the protons from water and represent the dissociation as follows:

$$CH_3COOH \rightleftharpoons CH_3COO^- + H^+$$

Now we can approximate $[H^+] = [CH_3COO^-] = x$.

In addition, because acetic acid is a weak acid, the concentration of the undissociated acid will not be appreciably lowered by dissociation. We can assume then that $[CH_3COOH] = 0.1\ M$.

Thus
$$K_a = \frac{[CH_3COO^-][H^+]}{[CH_3COOH]} = \frac{x^2}{0.1}$$
$$x^2 = 0.1 \times K_a = 0.1 \times 10^{-4.7}$$

Therefore $x = 0.0014\ \text{mol L}^{-1}$ and $pH = 2.85$

Note: The second assumption made above could be tested by setting $[CH_3COOH] = 0.1 - x$ and solving the quadratic equation for x. In that case, pH = 2.80. This answer is not very different from that obtained employing the simplifying assumption. The assumption, therefore, was reasonable.

3.9. At what pH values would glutamate be a good buffer?

SOLUTION

By inspection of Fig. 3-3, it can be seen that the pH of a glutamate solution changes least rapidly with NaOH concentration in the vicinity of the pK_a values, i.e., near pH 2.2, 4.3, and 9.7.

3.10. What amino acids would buffer best at pH values of (*a*) 2.0; (*b*) 6.0; (*c*) 4.5; (*d*) 9? (Refer to Table 3.2.)

SOLUTION

(*a*) All amino acids have a carboxylic group whose pK_a is close to 2.0. So all amino acids would be good buffers in this pH range.

(*b*) Histidine has a side chain whose pK_a value is close to 6. Histidine would therefore be a good buffer at pH 6.0. Inspection of Fig. 3-2 shows that the simple amino acids are very poor buffers at this pH.

(*c*) Glutamate has a side chain whose pK_a value is close to 4.5, so glutamate would buffer in this region.

(*d*) All amino acids have an α-amino group whose pK_a value is near 9, so all would buffer well at this pH value.

3.11. Derive a general equation that describes the average net charge (Z) on a prototropic group in terms of the solution pH and the pK_a of the group.

SOLUTION

By rearranging the Henderson-Hasselbalch equation (*3.13*), we obtain:

$$\frac{[\text{acid}]}{[\text{conjugate base}]} = 10^{(pK_a - pH)}$$

If we define α as the fraction of the group in the acid form, then

$$\alpha = \frac{[\text{acid}]}{[\text{acid}] + [\text{conjugate base}]} = \frac{10^{(pK_a - pH)}}{1 + 10^{(pK_a - pH)}}$$
$$= \frac{1}{10^{(pH - pK_a)} + 1}$$

Therefore, if the group (such as the amino group) has a cationic acid form, α represents the fractional positive charge:

$$Z = +\alpha = \frac{+1}{1 + 10^{(\mathrm{pH}-\mathrm{p}K_a)}}$$

Note: When $\mathrm{pH} \ll \mathrm{p}K_a$, Z approaches $+1$.

On the other hand, for groups such as the carboxyl, with a neutral acid form and an anionic conjugate base, α represents the fraction uncharged. The fractional charge is then:

$$Z = -(1-\alpha) = \alpha - 1 = \frac{-1}{1 + 10^{(\mathrm{p}K_a-\mathrm{pH})}}$$

In this case, when $\mathrm{pH} \ll \mathrm{p}K_a$, Z approaches zero.

In general, when the pH is more than 2 units below the $\mathrm{p}K_a$ value, the group can be considered completely protonated.

3.12. Using the $\mathrm{p}K_a$ values for aspartic acid listed in Table 3.2, determine the average net charge on the molecule at pH values of 1.0, 3.9, 6.8.

SOLUTION

From the solution to Prob. 3.11, the average net charge on each of the three groups can be determined at each of the pH values. The results are as follows:

Group	Average Charge		
	pH 1.0	pH 3.9	pH 6.8
α-COOH	−0.07	−0.98	−1.0
β-COOH	0	−0.5	−1.0
α-NH_3^+	+1.0	+1.0	+1.0
Average net charge	+0.93	−0.48	−1.0

3.13. In most amino acids, the $\mathrm{p}K_a$ of the α-carboxyl is near 2.0, and that of the α-amino is near 9.0. However, in peptides, the α-carboxyl has a $\mathrm{p}K_a$ of 3.8, and the α-amino has a $\mathrm{p}K_a$ of 7.8. Can you explain the difference?

SOLUTION

In the free amino acids, the neighboring charges affect the $\mathrm{p}K_a$ of each group. The presence of a positive $-NH_3^+$ group stabilizes the charged $-COO^-$ group, making the carboxyl a stronger acid; conversely, the negative carboxylate stabilizes the $-NH_3^+$ group, making it a weaker acid and thus raising its $\mathrm{p}K_a$. When peptides are formed, the free α-amino and carboxylate are farther apart, and exert less mutual influence.

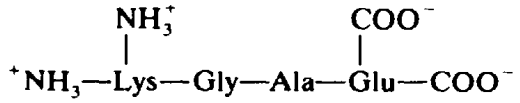
A secondary effect is that the carbonyl of the first peptide group has an inductive electron-withdrawing effect on the terminal α-amino, decreasing the stability of the $-NH_3^+$ and lowering the $\mathrm{p}K_a$ slightly.

3.14. Predict the migration direction (anodal, cathodal, or stationary) for the following peptide at pH 6.0: Lys-Gly-Ala-Glu.

SOLUTION

In peptides, the α-NH_3^+ and α-COO^- groups involved in peptide bond formation have been lost; only the N-terminal α-NH_3^+ and C-terminal α-COO^- groups remain. In addition, in the peptide above, lysine carries a positively charged side chain at pH 6.0, and glutamate carries a carboxylate group. The

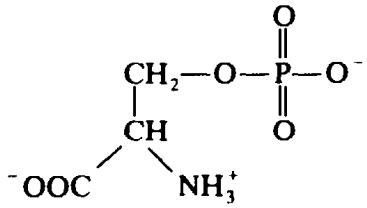
two amino groups, both on lysine, have pK_a values above 8. They will therefore both be nearly completely *protonated* and positively charged at pH 6.0. The two carboxyl groups on glutamate have pK_a values near 4. They will therefore be nearly totally *deprotonated* at pH 6.0. Thus, the peptide can be written



and the net charge is zero. The peptide will therefore be stationary.

AMINO ACID ANALYSIS

3.15. Serine phosphate



can be found in enzymatic hydrolysates of casein, a protein found in milk, yet this is not one of the 20 amino acids coded for in protein synthesis. Explain.

SOLUTION

In the synthesis of casein, serine is incorporated into the protein sequence. Subsequently, some of these serine residues are phosphorylated to form serine phosphate.

3.16. The elution profile shown in Fig. 3-4 indicates that glycine, alanine, valine, and leucine can be separated clearly by ion-exchange chromatography; yet these four amino acids have almost indistinguishable sets of pK_a values (Table 3.2). How can you account for this behavior?

SOLUTION

These amino acids all differ in their polarity. The polystyrene matrix that carries the charged groups of the ion-exchange resin is relatively nonpolar, and some separation will be achieved by means of partition effects, as well as by means of ion exchange. Amino acids such as leucine interact more strongly with the resin than do more polar amino acids such as serine, and this interaction retards passage through the column.

3.17. A small peptide was subjected to hydrolysis and amino acid analysis. In addition, because acid hydrolysis destroys tryptophan, the tryptophan content was estimated spectrophotometrically. From the following data, determine the empirical formula of the peptide.

Amino Acid	μmol
Ala	2.74
Glu	1.41
Leu	0.69
Lys	2.81
Arg	0.72
Trp	0.65

SOLUTION

Taking leucine as a convenient reference, we determine the molar ratios of the amino acids, which are 3.97, 2.04, 1.00, 4.07, 1.04, 0.94. Rounding these numbers off to integers yields the formula: Ala_4, Glu_2, Leu, Lys_4, Arg, Trp.

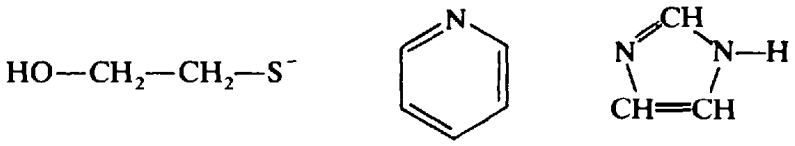
Supplementary Problems

3.18. List the amino acids with side chains capable of acting as (*a*) H-bond donors; (*b*) H-bond acceptors.

3.19. List those amino acids that absorb ultraviolet light strongly.

3.20. Which amino acid found in proteins does not show optical activity?

3.21. Write the conjugate acids of the following weak bases:



(*a*) (*b*) (*c*)

3.22. Calculate the isoelectric point of arginine.

3.23. Calculate the isoelectric point of histidine.

3.24. Calculate the pH of 1 *M* acetic acid.

3.25. What is the average net charge on lysine at the following pH values: (*a*) 2.0; (*b*) 5.0; (*c*) 10.0?

3.26. Using the information given in Prob. 3.13, determine the average net charge on the peptide *glycylglycylglycine* at pH: (*a*) 2.0; (*b*) 5.0; (*c*) 9.0.

Chapter 4

Proteins

4.1 INTRODUCTION

Proteins are naturally occurring polypeptides of molecular weight greater than 5,000. These *macromolecules* show great diversity in physical properties, ranging from water-soluble enzymes to the insoluble keratin of hair and horn, and they perform a wide range of biological functions.

What Biological Functions do Proteins Perform?

Proteins fulfill the following biochemical roles:

1. *Enzymatic catalysis*. Enzymes are protein catalysts, capable of enhancing rates of reactions by factors of up to 10^{12}.
2. *Transport and storage*. Many small molecules and ions are transported in the blood and within cells by being bound to *carrier proteins*. The best example is the oxygen-carrying protein *hemoglobin*. Iron is stored in various tissues by the protein *ferritin*.
3. *Mechanical functions*. Proteins often fulfill structural roles. The protein *collagen* provides tensile strength in skin, teeth, and bone. The membranes surrounding cells and cell organelles are also partly composed of proteins, having both functional and structural roles.
4. *Movement*. Muscle contraction is accomplished by the interaction between two types of protein filaments, *actin* and *myosin*. Myosin also possesses an enzymatic activity for facilitating the conversion of the chemical energy of ATP into mechanical energy.
5. *Protection*. The antibodies are proteins, aided in mammals by *complement*, a complex set of proteins involved in the destruction of foreign cells.
6. *Information processing*. Stimuli external to a cell, such as hormone signals or light intensity, are detected by specific proteins that transfer a signal to the interior of the cell. A well-characterized example is the visual protein *rhodopsin*, located in membranes of retinal cells.

Question: What is a common feature among those functions listed above that may be explained in terms of protein structure?

In all the above examples, the phenomenon of *specific binding* is involved. For example, hemoglobin specifically binds molecular oxygen, antibodies bind to specific foreign molecules, enzymes bind to specific substrate molecules and in doing so bring about selective chemical bond rearrangement. The function of a protein, then, is understood in light of how the structure of the protein allows the specific binding of particular molecules.

4.2 PURIFICATION AND CHARACTERIZATION OF PROTEINS

Purification

The first step in protein purification often involves separating the protein molecules from low-molecular-weight solutes. Some degree of separation of different proteins can then be achieved

on the basis of physical properties such as electrical charge, molecular size, and differential solubility in various solvents. Finally, high purification can be attained on the basis of specific affinity for certain compounds that are linked to some form of solid support, the process known as *affinity chromatography*.

Dialysis

Protein molecules are of high molecular weight, by definition greater than 5,000. Thus, they will not pass through a cellophane membrane that allows free passage to smaller molecules. This selective permeability is the basis of the process termed *dialysis*, which is usually carried out by placing the protein solution inside a cellophane bag and immersing the bag in a large volume of buffer. The small molecules diffuse across the bag into the surrounding buffer, while the proteins remain within the bag. Dialysis tubing can be obtained with molecular weight cut-off limits ranging from 1,000 to more than 10,000.

Selective Solubility

Proteins may be selectively precipitated from a mixture of different proteins by adding (1) neutral salts such as ammonium sulfate (*salting out*), (2) organic solvents such as ethanol or acetone, or (3) potent precipitating agents such as trichloroacetic acid. Proteins are *least* soluble in any given solvent when the pH value is equal to their *isoelectric point* (pH_I). At the isoelectric point, the protein carries no net electric charge, and therefore electrostatic repulsion is minimal between protein molecules. Although a protein may carry both positively and negatively charged groups at its isoelectric point, the *sum* of these charges is zero. At pH values *above* the isoelectric point, the net charge will be *negative*, while at pH values *below* the pH_I, the net charge will be *positive*.

Electrophoresis

The movement of electrically charged protein molecules in an electrical field, termed *electrophoresis*, is an important means of separating different protein molecules. In electrophoresis, proteins are banded into narrow zones, providing that the effects of convection are minimized, for example, by the use of a stabilizing support such as paper or polyacrylamide gel.

The terms *anode* and *cathode* are often used in electrophoresis and are often confused. Remember that an *anion* is a negatively charged ion; *anions* move toward the *anode*.

EXAMPLE 4.1

In what direction will the following proteins move in an electric field [toward the anode, toward the cathode, or toward neither (i.e., remain stationary)]: (*a*) egg albumin ($pH_I = 4.6$) at pH 5.0; (*b*) β-lactoglobulin ($pH_I = 5.2$) at pH 5.0; at pH 7.0?

(*a*) For egg albumin, pH 5.0 is above its isoelectric point, and the protein will therefore carry a *small* excess negative charge. It will thus migrate toward the anode.

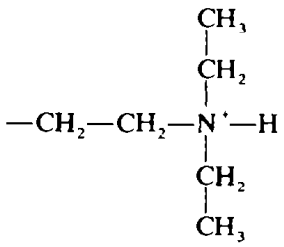
(*b*) For β-lactoglobulin, pH 5.0 is below its isoelectric point; at this pH, the protein will be positively charged and will move toward the cathode. At pH 7.0, on the other hand, the protein will be negatively charged and will move toward the anode.

Ion-Exchange Chromatography

Ion-exchange chromatography relies on the electrostatic interaction between a charged protein and a stationary ion-exchange-resin particle carrying a charge of opposite sign. The strength of attraction between the protein and the resin particle depends on the charge on the protein (and thus on the solution pH) and on the dielectric constant of the medium [Eq. (*4.2*)]. The interaction can be modified in practice by altering the pH or the salt concentration.

EXAMPLE 4.2

A solution containing egg albumin ($pH_I = 4.6$), β-lactoglobulin ($pH_I = 5.2$), and chymotrypsinogen ($pH_I = 9.5$) was loaded onto a column of diethylaminoethyl cellulose (DEAE-cellulose) at pH 5.4. The column was then *eluted* with pH 5.4 buffer, with an increasing salt concentration. Predict the elution pattern.



Diethylaminoethyl (DEAE) group

DEAE-cellulose carries a positive charge at pH 5.4. At this pH value, chymotrypsinogen also carries a positive charge, and therefore does not bind. β-Lactoglobulin, carrying a very small negative charge, binds only weakly and will be displaced as the salt concentration is raised. However, egg albumin, carrying a larger negative charge, binds tightly and will be displaced only at high salt concentrations, or on lowering the pH to a value at which its negative charge is reduced.

Gel Filtration

Separation of proteins on the basis of size can be achieved by means of *gel filtration*. This technique relies on diffusion of protein molecules into the pores of a *gel matrix* in a column. A commonly used type of gel material is *dextran*, a polymer of glucose, in the form of very small beads. This material is available commercially as *Sephadex* in a range of different pore sizes.

When the protein is larger than the largest pore of the matrix, no diffusion into the matrix takes place and the protein is eluted rapidly. Molecules smaller than the smallest pore can diffuse freely into all gel particles and elute later from the column, after a larger volume of buffer has passed through.

EXAMPLE 4.3

In what order would the following proteins emerge upon gel filtration of a mixture on Sephadex G-200: myoglobin ($M_r = 16{,}000$), catalase ($M_r = 500{,}000$), cytochrome *c* ($M_r = 12{,}000$), chymotrypsinogen ($M_r = 26{,}000$), and serum albumin ($M_r = 65{,}000$)?

Catalase is the largest protein in this set and would be excluded completely from the Sephadex beads (exclusion limit approx. 200,000). That means that to the catalase molecules, the beads appear to be solid, and so the catalase molecules can only dissolve in the fluid outside the beads (the *void volume*). They would be eluted from the column when a volume of eluent equal to this void volume had passed through the column.

Cytochrome *c* is the smallest protein of the set and would be able to diffuse freely into all the space within the beads. The column would therefore appear to have a larger volume available to cytochrome *c*, which would not emerge until the column had been flushed with a volume almost equal to its total volume.

The order of elution would therefore be: catalase, serum albumin, chymotrypsinogen, myoglobin, and cytochrome *c*.

Sequencing Proteins

The structure and properties of peptides and proteins depend critically upon the sequence of amino acids in the peptide chain. The first complete amino acid sequence of a protein, that of insulin (51 amino acid residues), was determined by F. Sanger in 1953. The process is now performed using automated protein sequencers, and involves step-by-step identification of amino acids at the N terminus of the protein using a chemical process known as *Edman degradation*.

N-Terminal Sequencing

The N-terminal residue, i.e., the first amino acid in the sequence of a peptide, can be determined by reaction with *phenylisothiocyanate*. At neutral pH, this compound reacts with the α-amino group. After mild acid hydrolysis, the reaction product cyclizes, releasing the terminal residue as a *phenylthiohydantoin* (PTH) derivative (the Edman degradation, Fig. 4-1). The derivative can be analyzed to determine its parent amino acid and its quantity.

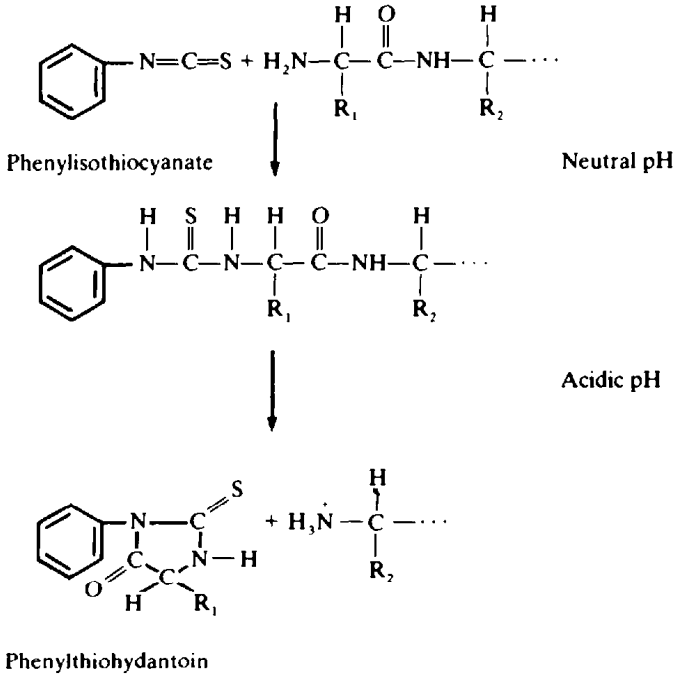
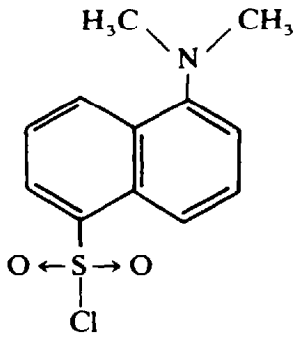
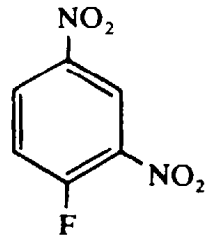


Fig. 4-1 The Edman degradation.

Other reagents capable of reacting with the amino-terminal residues, such as *dansyl chloride* and *fluoro-2,4-dinitrobenzene*, may also be used to identify the N terminus.



Dansyl chloride



Fluoro-2,4-dinitrobenzene

EXAMPLE 4.4

Given a means of identifying and quantifying the N-terminal residue, describe a way of determining the entire sequence of a protein.

If the phenylisothiocyanate method is used, the cyclization and release of the N-terminal derivative occurs under mild conditions that leave the rest of the chain intact. It is therefore possible to take the protein chain, now without its original N-terminal residue, and repeat the procedure to determine the second residue in the sequence, and so on. Unfortunately, at each step, there is a finite chance of additional peptide hydrolysis or incomplete reaction, and uncertainty tends to accumulate after 10 to 20 cycles.

Specific Cleavage of Peptides

Many proteins contain hundreds of amino acid residues, and it may not be feasible to determine the sequence of the whole molecule in one go, because of the buildup of uncertainty at each step. It is therefore convenient first to cleave the protein into smaller, more manageable pieces. Also the protein may contain *disulfide bridges* linking cysteine residues in different parts of the chain; these links must be broken to allow sequencing to continue.

Trypsin

One of the most frequently used means of breaking polypeptides into smaller fragments is hydrolysis by the enzyme *trypsin*. This digestive enzyme, produced by the pancreas, hydrolyzes peptide bonds on the carboxyl side of lysine and arginine (i.e., positively charged) residues:

$$R_1\text{-Lys-Ala-}R_2 \xrightarrow{\text{Trypsin}} R_1\text{-Lys-COO}^- + NH_3^+\text{-Ala-}R_2$$

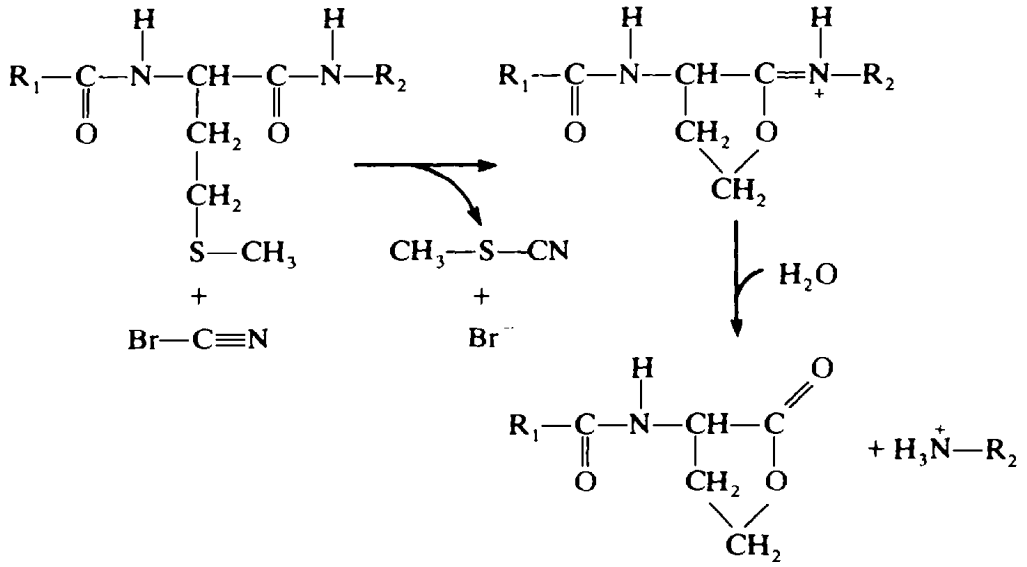
Chymotrypsin

Another commonly used enzyme for selective cleavage of peptides is *chymotrypsin*, which hydrolyzes peptide bonds on the carboxyl side of aromatic residues (phenylalanine, tyrosine, and tryptophan):

$$R_1\text{-Phe-Ser-}R_2 \xrightarrow{\text{Chymotrypsin}} R_1\text{-Phe-COO}^- + NH_3^+\text{-Ser-}R_2$$

Cyanogen Bromide Cleavage

A nonenzymatic, chemical method for specific cleavage of polypeptides involves the reaction of *cyanogen bromide*, CNBr, with methionine residues:



Overlapping Peptides

When a protein is cleaved into manageable peptides and the sequence of each peptide is determined, the next problem is that of ordering the peptides themselves in the correct sequence. For this, at least two sets of peptide sequences from different selective cleavage methods are required.

EXAMPLE 4.5

A peptide was broken into two smaller peptides by cyanogen bromide (CNBr), and into two different peptides by trypsin (Tryp). Their sequences were as follows: CNBr 1, Gly-Thr-Lys-Ala-Glu; CNBr 2, Ser-Met; Tryp 1, Ser-Met-Gly-Thr-Lys; Tryp 2, Ala-Glu. Determine the sequence of the parent peptide.

By arranging the sequences in an overlapping set, as follows, it is possible to determine the parent sequence:

```
CNBr 2:  Ser-Met
Tryp 1:  Ser-Met-Gly-Thr-Lys-
CNBr 1:          Gly-Thr-Lys-Ala-Glu
Tryp 2:                    Ala-Glu
         ---------------------------
         Ser-Met-Gly-Thr-Lys-Ala-Glu
```

Note: Some of this information is redundant; peptides Tryp 1 and CNBr 1 would be sufficient for unambiguous sequence determination. In addition, in such a simple case, with only two peptides from each cleavage reaction, we could determine the sequence from either one of these cleavages alone using the fact that the carboxyl sides of Met and Lys are the sites of cleavage with CNBr and trypsin, respectively.

Some proteins also contain other compounds, apart from amino acids. These proteins are known as *conjugated proteins* and the non-amino acid part is known as a *prosthetic group* (Chap. 8); the protein part is termed the *apoprotein*. Glycoproteins (Chap. 6) and proteoglycans (Chap. 5) contain covalently bound carbohydrate, while *lipoproteins* (Chap. 6) contain lipid as the prosthetic groups.

Molecular Weight Determination

Each protein has a unique *molecular weight*. Furthermore, the size or molecular weight of a particular protein, under specified conditions, distinguishes it from many other proteins. Note that the molecular weight (or more precisely, the *relative molecular weight*), M_r, is dimensionless, and represents the mass of the molecule relative to 1/12 of the mass of an atom of ^{12}C. Molecular *mass*, on the other hand, is usually expressed in units of daltons (Da) or kilodaltons (kDa), where 1 Da is 1/12 of the mass of an atom of ^{12}C (1.66×10^{-24} g). The *molar mass* is the mass of one mole expressed in grams. All three quantities have the same *numerical* value, but have different units. For example, serum albumin could be described as having a molecular weight of 66,000, a molecular mass of 66 kDa, or a molar mass of 66 kg mol^{-1}.

The molecular weight is determined by the amino acid sequence of the protein, any post-translational modifications such as glycosylation, and the presence of bound nonpeptide groups such as metals and cofactors. While the exact weight of a polypeptide chain can be calculated from its amino acid sequence, a *rule-of-thumb* calculation, based on the proportion of different amino acids found in most proteins, gives the molecular weight as the number of residues multiplied by 110.

EXAMPLE 4.6

Human cytochrome *c* contains 104 amino acid residues. What is its approximate molecular weight?

The molecular weight is approximately 11,900, the sum of the polypeptide molecular weight (approx. 104×110) and that of the heme prosthetic group (412). Alternatively, the molecular mass is approx. 11.9 kDa.

Proteins that possess a *quaternary* structure are composed of several separate polypeptide chains held together by noncovalent interactions. When such proteins are examined under dissociating conditions (e.g., 8 *M* urea to weaken hydrogen bonds and hydrophobic interaction, 1 m*M* mercaptoethanol to disrupt disulfide bonds), the molecular weight of the component polypeptide chains can be determined. By comparison with the *native* molecular weight, it is often possible to determine how many polypeptide chains are involved in the native structure.

The molecular weight of a protein may be determined by the use of thermodynamic methods, such as *osmotic pressure* measurement, and *sedimentation* analysis in the ultracentrifuge. Osmotic

pressure is sensitive to the *number* of molecules in solution, and if one knows the total mass of protein in solution, the molar weight can be calculated. The sedimentation of a protein molecule in a centrifuge depends directly on the *mass* of the molecule.

It is also possible to estimate the molecular weight of a protein by means of gel filtration, or from measurement of its rate of migration through an electrophoresis gel (*gel electrophoresis*), in comparison with a series of standards of known molecular weight. An absolute and very accurate means of determining protein molecular weights is *mass spectrometry*.

Sedimentation Analysis

The molar mass (M) of a solute can be determined from measurements of the *sedimentation coefficient*, s, and the *diffusion coefficient*, D (see Prob. 4.18), according to the Svedberg equation:

$$M = \frac{RTs}{D(1 - \bar{v}\rho)} \tag{4.1}$$

where R is the gas constant, T the absolute temperature, $\bar{v}$ the *partial specific volume* of the solute (the reciprocal of its density), and ρ the solvent density. The sedimentation coefficient, s, which depends on the molar weight of the solute, as well as its size and shape, is determined from the rate at which the solute sediments in the gravitational field of an ultracentrifuge. The diffusion coefficient, D, is dependent on size and shape and is determined from the rate of spreading of the boundary between a solution and the pure solvent.

The sedimentation coefficient is often reported in Svedberg units, S, named in honor of Thé Svedberg, the inventor of the ultracentrifuge: $1\ \text{S} \equiv 10^{-13}$ seconds. Sedimentation coefficients of some selected proteins are listed below.

Protein	M_r	s
Lysozyme	14,000	1.9 S
Hemoglobin	64,500	4.5 S
Catalase	248,000	11.3 S
Urease	480,000	18.6 S

Gel Filtration

The volume required to elute a globular protein from a gel-filtration column is a monotonically decreasing function of molecular weight. Thus, by comparing the elution behavior of an unknown protein with that of a series of standards of known molecular weight, an estimation of molecular weight can be interpolated (Fig. 4-2).

EXAMPLE 4.7

A series of standard proteins and an unknown enzyme were examined by means of gel filtration on Sephadex G-200. The elution volume, V_{el}, for each protein is given below.

(*a*) Plot the data in the form of log M_r versus elution volume. (*b*) From the line of best fit through the points for the standards, determine the molecular weight of the unknown enzyme. (*c*) Explain why ferritin and ovomucoid behave anomalously.

(*a*) Figure 4-2 shows the line of best fit for the data and indicates coordinates for the unknown.

(*b*) From Fig. 4-2, M_r of the unknown enzyme $\simeq$ 126,000.

(*c*) Ferritin contains an iron hydroxide core, and therefore its density is higher than that of the standards. Ovomucoid is a glycoprotein of different density, and probably different shape, from the standards.

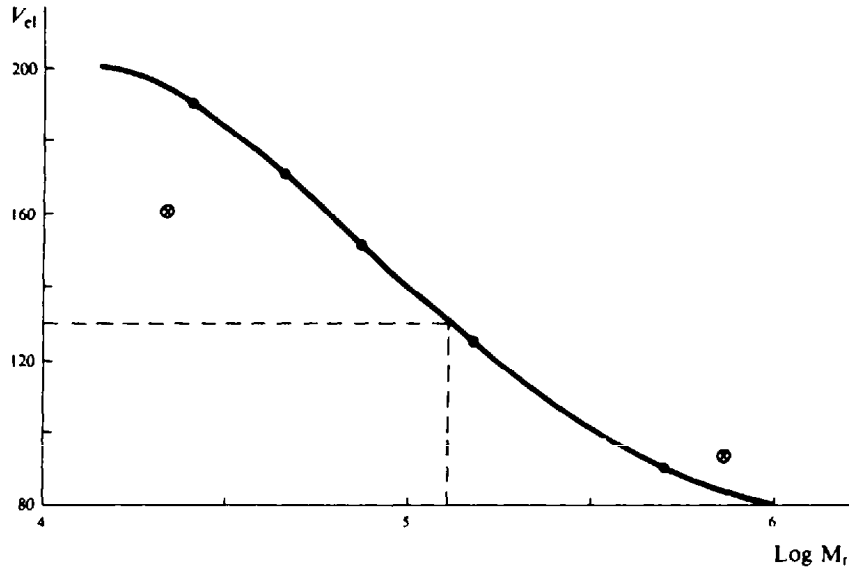


Fig. 4-2 Gel filtration of proteins. Elution volumes V_{el} are plotted for a series of proteins of different M_r applied to a column of Sephadex G-200. The two data points denoted by ⊗ were not used in constructing the calibration curve.

Protein	M_r	V_{el} (mL)
Blue dextran*	10^6	85
Lysozyme	14,000	200
Chymotrypsinogen	25,000	190
Ovalbumin	45,000	170
Serum albumin	65,000	150
Aldolase	150,000	125
Urease	500,000	90
Ferritin†	700,000	92
Ovomucoid†	28,000	160
Unknown	——	130

*Blue dextran is a high molecular-weight carbohydrate with a covalently bound dye molecule.
†Do not use for calibration.

Gel Electrophoresis

In the presence of the detergent *sodium dodecyl sulfate* (SDS), many proteins are disaggregated and unfolded. As with elution volume, the electrophoretic mobility of SDS-denatured polypeptide chains is a monotonically decreasing function of molecular weight, and the molecular weight of an unknown protein may be inferred from the mobilities of standards.

Question: What assumptions are inherent in both of the above techniques for the determination of molecular weight?

1. In both cases, the hydrodynamic behavior depends on the *shape* of the molecule. We must, therefore, assume that the shape (i.e., spherical, ellipsoidal, etc.) of the unknown protein is the same as that of the standards.

2. The gel-filtration behavior really depends on the *effective size*, not mass. Therefore, if the protein differs from the standards in density, an incorrect estimate of molecular weight will result.
3. Proteins bind SDS, and as a first approximation the amount of SDS bound per gram is the same for all proteins. Differences in the amount bound per molecule will result in differences in the total charge, leading to differences in the electrophoretic mobility, and an incorrect value of molecular weight will be inferred.

Mass Spectrometry

Previous use of mass spectrometry to determine the molecular weights of biological macromolecules was limited by the lack of a means of vaporizing the sample without causing thermal decomposition. In recent years a number of techniques have been developed to overcome these problems and mass spectrometry is becoming a routine tool in most biochemistry laboratories. The methods which are now used to introduce the sample include: *matrix assisted laser desorption* (MALDI), in which the sample is embedded in a low molecular weight organic matrix and is irradiated with a pulsed UV laser; *electrospray*, in which a dilute acidic solution of the protein sample is sprayed from a positively charged needle into the vacuum chamber (the solvent surrounding the protein is rapidly evaporated leaving the bare protein molecules carrying multiple positive charges); and *plasma desorption*, which uses high energy charged particles to liberate the sample from a metal support (the positively charged particles are accelerated by an electric field and then deflected by a magnetic field, where they are separated by their mass-to-charge ratio). Mass spectrometry is accurate to about 0.01%, requires only picomoles of sample, and can be used to determine the molecular weights of molecules ranging from single amino acids to proteins larger than 100 kDa.

4.3 PROTEIN FOLDING

Proteins Generally Fold into Compact, Well-Defined, Three-Dimensional Structures

In order to understand the functions of proteins, we need to know something about the *conformation*, or the three-dimensional folding pattern, that the polypeptide chain adopts. Although many artificial polyamino acids have no well-defined conformation and seem to exist in solution as nearly random coils, most biological proteins adopt a well-defined *folded structure*. Some, such as the keratins of hair and feathers, are *fibrous* and organized into linear or sheetlike structures with a regular, repeating folding pattern. Others, such as most enzymes, are folded into compact, nearly spherical, *globular* conformations.

Question: Why do proteins fold?

The folding of a protein into a compact structure is accompanied by a large *decrease in conformational entropy* (disorder) of the protein, which is thermodynamically unfavorable. The native, folded conformation is maintained by a large number of *weak*, *noncovalent interactions* that act cooperatively to offset the unfavorable reduction in entropy. These *noncovalent interactions* include hydrogen bonds, and electrostatic, hydrophobic, and van der Waals interactions. These interactions ensure that the folded protein is (often just marginally) more stable than the unfolded form.

Electrostatic Interactions

Charged particles interact with one another according to Coulomb's law:

$$\Delta E = \frac{Z_A Z_B \varepsilon^2}{D r_{AB}} \tag{4.2}$$

where ΔE is the energy of the electrostatic interaction, Z_A and Z_B are the number of charges on the two interacting species, r_{AB} is the distance between the two species, ε is the charge of an electron, and D is the dielectric constant of the medium. When the two charges are of opposite sign, the interaction energy decreases as they approach each other, and hence the interaction is favorable.

Thus, negatively charged moieties in proteins (such as the carboxylate side chains of Asp and Glu residues) frequently interact with positively charged side chains of Lys, Arg, or His residues. These electrostatic interactions often result in the formation of *salt bridges*, in which there is some degree of hydrogen bonding in addition to the electrostatic attraction, as illustrated in Fig. 4-3.

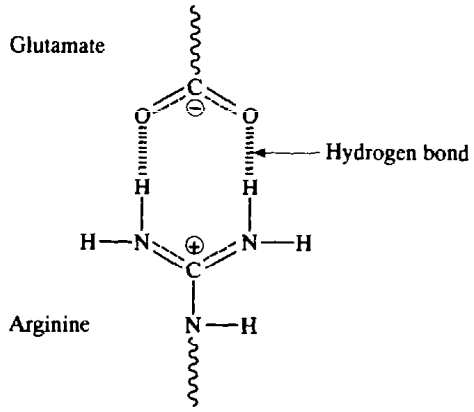


Fig. 4-3 A salt bridge between the side chains of an Arg and Glu residue.

Since water has a high dielectric constant of 80, the energy associated with an ion pair in a protein ranges from as low as 0.5–1.5 kJ mol^{-1} for a surface interaction up to 15 kJ mol^{-1} for an electrostatic interaction between residues buried in the interior of the protein, where the dielectric constant is expected to be lower.

van der Waals Interactions

All atoms and molecules attract one another as a result of transient dipole–dipole interactions. A molecule need not have a net charge to participate in a dipolar interaction; electron density can be highly asymmetric if interacting atoms have different electronegativities.

Atoms with the greatest electronegativities have the largest "excess" of negative charge: the electronegativities of the atoms found in proteins are O, 3.44; N, 3.04; C, 2.55; and H, 2.20 (on a scale of 0.8–4).

These transient dipolar interactions are known as van der Waals interactions. They are weak and close-range, varying inversely as the sixth power of the interatomic distance. When atoms involved in a van der Waals interaction approach too closely, there is a strong repulsive interaction. Thus, the van der Waals interaction energy (E_{vdw}) is usually given by the following equation:

$$E_{vdw} = \frac{a}{d^{12}} - \frac{b}{d^6} \tag{4.3}$$

where d is the interatomic distance, and a and b are positive constants. The first and second terms in the equation represent the repulsive and attractive parts of the van der Waals interaction, respectively. Eq. (*4.3*) is often referred to as the *Lennard-Jones 6,12 potential.*

The optimal distance for a van der Waals interaction is when the interacting atoms are separated by 0.3–0.5 Å more than the sum of their van der Waals radii (defined as the minimum contact distance observed between the atoms in a crystal). The van der Waals interaction energy is rather small, usually less than 1 kJ mol^{-1}.

Hydrogen Bonds

A hydrogen bond results from an electrostatic interaction between a hydrogen atom covalently bound to an *electronegative* atom (such as O, N or S), and a second electronegative atom with a *lone pair* of *nonbonded* electrons:

$$\underset{\text{Donor}}{-O-H} \text{ ------ } \underset{\text{Acceptor}}{O=C}$$

Although the hydrogen atom is formally bonded to the *donor* group, it is partly shared between the *donor* and the *acceptor*. Hydrogen bonds are highly directional and are strongest when all three participating atoms lie in a straight line. Most hydrogen bonds in proteins occur between the backbone C=O and N—H groups, with an H···O distance of 1.9–2.0 Å. It has been estimated that an average hydrogen bond of this type contributes ~5 kJ mol^{-1} to the stability of a protein in aqueous solution, although this value may vary from 2 kJ mol^{-1} to 7.5 kJ mol^{-1} (see Table 4.1).

Table 4.1. Types of Noncovalent Interactions Involved in Stabilizing Protein Structure

Interaction	Example	Bond energy (kJ mol^{-1})*
van der Waals	C—H···H—C	0.4–2.0
Electrostatic	$-COO^- \cdots H_3N^+-$	0.5–15
Hydrogen bond	—N—H···O=C—	2.0–7.5
Hydrophobic†	Burial of $-CH_2-$	~3

*The bond energy is the energy to break the interaction.
†This value represents the free energy required to transfer a $-CH_2-$ group of a nonpolar side chain from a protein's interior to water.

Hydrophobic Interactions

The placing of a nonpolar group in water leads to energetically unfavorable ordering of the water molecules around it, i.e., a lowering of the entropy of the solution. Transfer of nonpolar groups from water to a nonpolar environment is thus accompanied by an increase in entropy (of the water molecules) and is spontaneous (Chap. 10). The folding of a protein chain into a compact globular conformation removes nonpolar groups from contact with water; the increase in entropy arising from the liberation of water molecules compensates for the decrease in entropy of the folded polypeptide chain. Burial of a methylene ($-CH_2-$) group in the interior of a protein is as energetically favorable (~3 kJ mol^{-1}) as a strong hydrogen bond.

Hydrophobic interactions are such an important driving force in the folding of water-soluble globular proteins that we can formulate the general rule: *hydrophobic residues tend to be buried in the interior of proteins which minimizes their exposure to water.*

EXAMPLE 4.8

If the overall folding energy of a particular protein is only 40 kJ mol^{-1}, how many H bonds would have to be broken in order to disrupt the structure?

Since each H bond contributes an average of ~5 kJ mol^{-1} of stabilizing energy, the breaking of about eight such bonds would be sufficient to disrupt the native structure.

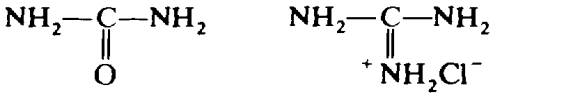
Protein Denaturants

Because the stabilization energy of most proteins is so small, many proteins show rapid, small fluctuations in structure, even at normal temperatures. In addition, it is fairly easy to cause protein molecules to unfold, or *denature*. Common denaturation agents are:

(*a*) High temperature

(*b*) Extremes of pH

(*c*) High concentrations of compounds such as urea or guanidine hydrochloride:



Urea Guanidine hydrochloride

(*d*) Solutions of detergents such as sodium dodecyl sulfate, $CH_3(CH_2)_{10}CH_2OSO_3^-Na^+$

Protein Structure is Dictated by Amino Acid Sequence

Several proteins and enzymes, completely unfolded by urea and with disulfide bridges reduced, are capable of refolding to the active, native state on removal of urea. This demonstrates that the *information* for the correct folding pattern exists in the sequence of amino acids.

EXAMPLE 4.9

Ribonuclease, the enzyme that hydrolyzes *ribonucleic acids* (Chap. 7), contains four disulfide bonds that help to stabilize its conformation. In the presence of 6 *M* guanidine hydrochloride, to weaken hydrogen bonds and hydrophobic interactions, and 1 m*M* mercaptoethanol, to reduce the disulfide bonds, all enzymatic activity is lost, and there is no sign of residual secondary structure. On removing the guanidine hydrochloride by dialysis or gel filtration, enzymatic activity is restored, the native conformation is regained, and correct disulfide bonds are reformed.

Many larger proteins require help in the form of *chaperonins* (which are themselves protein molecules) in order to fold correctly. These chaperonins are thought to act by trapping incorrectly folded intermediates, causing them to unfold and have another chance to fold correctly.

4.4 PROTEIN STRUCTURE

How Can the Structure of a Protein be Described?

One could completely describe the three-dimensional structure of a protein by placing it in a Cartesian coordinate system and listing (x, y, z) coordinates for each atom in the protein. Indeed, this is how experimentally determined protein structures are stored in the Protein Data Bank at the Brookhaven National Laboratory in the United States of America. However, the structure of a protein can be more succinctly described by listing the angles of rotation (*torsion angles*) of each of the bonds in the protein (see Fig. 4-4). For example, the backbone conformation of an amino acid residue can be specified by listing the torsion angles ϕ (rotation around the N—C_α bond), ψ (rotation around the C_α—C′ bond), and ω (rotation around the N—C′ bond).

The zero position for ϕ is defined with the —N—H group *trans* to the C_α—C′ bond, and for ψ with the C_α—N bond *trans* to the —C=O bond (Fig. 4-4). The peptide-bond torsion angle (ω) is generally 180° (see Example 4.10). A full description of the three-dimensional structure of a protein also requires a knowledge of the side-chain χ torsion angles.

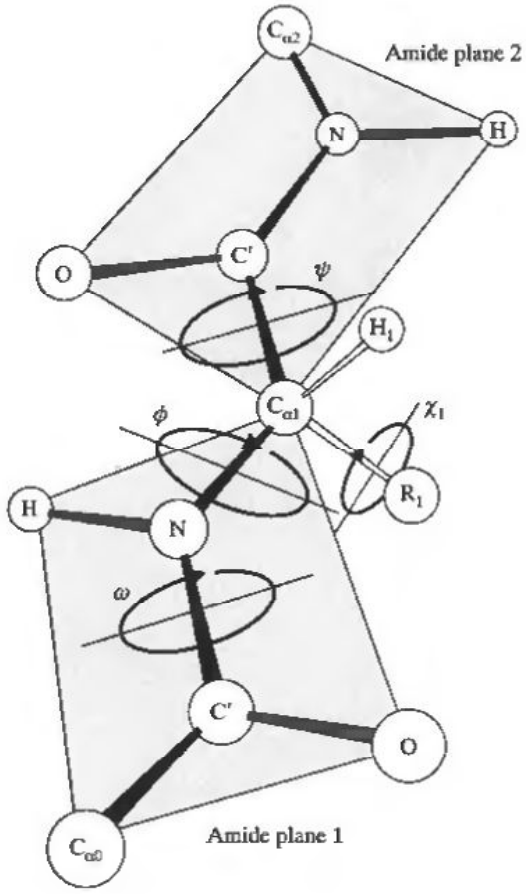
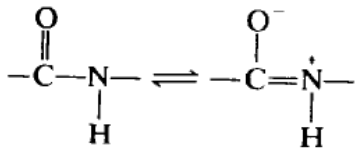


Fig. 4-4 Definition of protein torsion angles ω, ϕ, ψ, and χ. The conformation shown is obtained when ω, ϕ and ψ are all set to 180°.

EXAMPLE 4.10

Why is the peptide group planar?

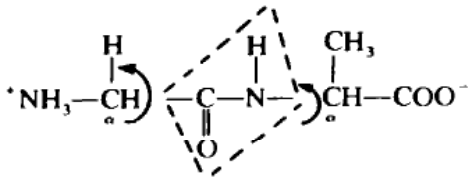
The C—N bond has a partial double-bond character owing to resonance between the two forms shown below:



The length of the C—N bond (0.132 nm) is intermediate between that of a C—N single bond (0.149 nm) and a C=N double bond (0.129 nm). The partial double-bond character restricts rotation around the C—N bond, such that the favored arrangement is for the O, C, N, and H atoms to lie in a plane, with the O and H atoms *trans*; this corresponds to a torsion angle of $\omega = 180°$. The *cis* conformation, which generally only occurs in proteins at X–Pro peptide bonds, corresponds to $\omega = 0°$.

EXAMPLE 4.11

In the dipeptide *glycylalanine*, which bonds of the backbone allow free rotation of the attached groups?



Glycylalanine

In this structure, the peptide group is indicated by the dashed lines. Because the peptide group itself is rigid and planar, there is no rotation around the bond between the carbonyl carbon atom and the nitrogen atom (the C′—N bond). However, free rotation is possible around the bond between the α carbon and the carbonyl carbon atom (the C_α—C′ bond) and about the bond between the nitrogen atom and the alanyl α-carbon atom the (N—C_α bond). Thus, for every peptide group in a protein, there are two rotatable bonds, the relative angles of which define a particular backbone conformation.

Question: Are there any restrictions on the structures that protein molecules can form?

Not all combinations of ϕ and ψ angles are possible, as many lead to clashes between atoms in adjacent residues. For all residues except glycine, the existence of such steric restriction involving side-chain atoms reduces drastically the number of possible conformations. The possible combinations of ϕ and ψ angles that do not lead to clashes can be plotted on a conformation map (also known as a *Ramachandran plot*, named after the chemist who did much of the pioneering work in this field). Figure 4-5 shows a Ramachandran plot for the allowed conformations of alanylalanine. The

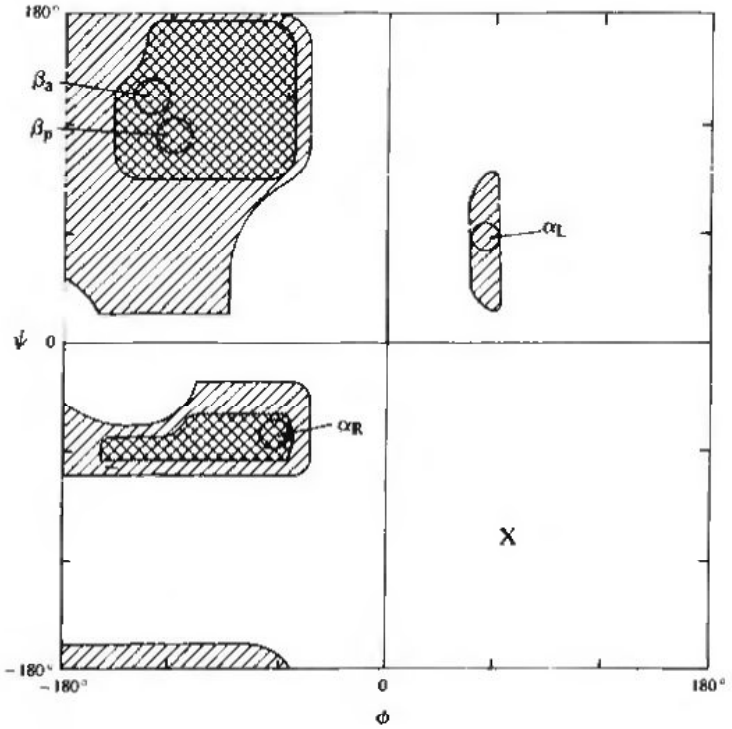


Fig. 4-5 Ramachandran plot for alanylalanine, showing the fully allowed regions (double-hatched) and partially allowed regions (single-hatched) of ϕ and ψ angles (see Fig. 4-4). The coordinates for the parallel and antiparallel β structures (β_p and β_a, respectively) and for the left-handed and right-handed α helices (α_L and α_R, respectively) are indicated.

double-hatched areas represent conformations (combinations of ϕ and ψ) for which no hindrance exists. The single-hatched areas represent conformations for which some hindrance exists, but which may be possible if the distortion can be compensated for by interactions elsewhere in the protein.

EXAMPLE 4.12

The right-handed α helix has ϕ and ψ values of $-57°$ and $-47°$, respectively. Would you expect an α helix to be a stable structure?

Yes. These values put the right-handed α helix into a particularly favorable area of the Ramachandran plot, indicated in Fig. 4-5 with the symbol α_R.

Proteins Often Contain Regular Repeating Structures

If the backbone torsion angles of a polypeptide are kept constant from one residue to the next, a regular repeating structure will result. While all possible structures of this type might be considered to be helices from a mathematical viewpoint, they are more commonly described by their appearance; hence α helices and β pleated sheets.

Given the normal van der Waals radii of the atoms, expected bond angles, and the planarity of the peptide bond, only two regular, repeating structures exist without distortion and with maximum hydrogen-bond formation:

1. The α helix, found in the α-keratins
2. The β pleated sheets (parallel and antiparallel), as exemplified by the β form of stretched keratin and silk protein

These regular repeating structures are also commonly found as elements of folding patterns in the globular proteins as well as the principal structure of fibrous proteins. For many globular proteins, a significant proportion of the polypeptide chain displays no regularity in folding. These regions may have short sections of commonly found structures such as *reverse turns*, which often link the strands of β sheets. Those regions without a regular repeating secondary structure are often referred to as having a *random-coil* conformation. However, these regions may still be well-defined, even if they are not regular, and they are better referred to as *disordered regions*.

The α Helix

In the α helix, the polypeptide *backbone* is folded in such a way that the $-C=O$ group of each amino acid residue is *hydrogen-bonded* to the $-N-H$ group of the *fourth* residue along the chain: i.e., the $-C=O$ group of the first residue bonds to the $-N-H$ group of the fifth residue, and so on.

Question: Do the $-N-H$ groups on residues 1 to 4 also hydrogen bond to other groups in the α helix?

There are no available $-C=O$ groups with which these amide groups can interact. Consequently, they remain unbonded. Similarly, at the other end of the chain, four $-C=O$ groups remain unbonded. If the polypeptide chain is very long, the lack of bonding at the ends has a negligible effect on the overall stability. However, short α-helical chains are less stable because the end effects are *relatively* more important.

The backbone of the α helix winds around the long axis, as shown in Fig. 4-6. The hydrogen bonds are all aligned approximately parallel to this axis, and the side chains protrude outward. Each residue is spaced 0.15 nm from the next along the axis, and 3.6 residues are required to make a

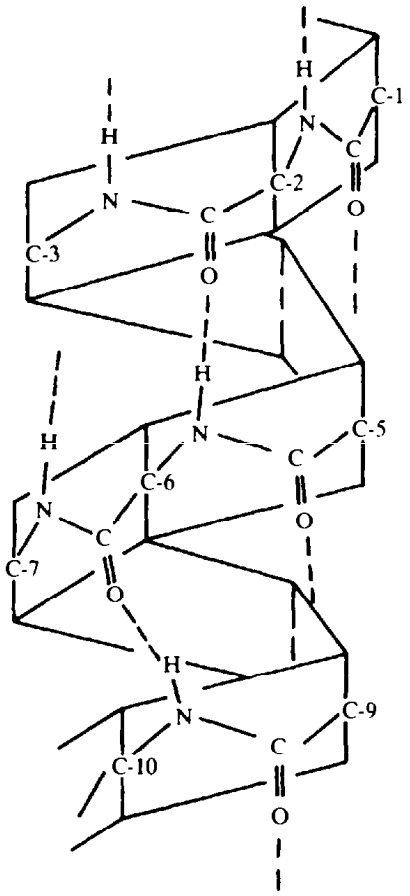


Fig. 4-6 The right-handed α helix. This diagram shows the peptide groups represented as planar segments, with the α -carbon atom at the junctions of successive planes.

complete turn of the helix. Although both left and right screw senses are possible, the right-hand screw sense is energetically favored with L-amino acids.

Because each $-C=O$ and $-N-H$ group is hydrogen-bonded (except for the four at each end), the α helix is strongly stabilized. However, for some amino acids, interactions involving the side chains may weaken the α helix, making this conformation less likely in polypeptide chains containing high proportions of such helix-destabilizing amino acids (Table 4.2).

Table 4.2. Tendency of Amino Acid Residues to Form α Helices

Helix formers	Glu, Ala, Leu, His, Met, Gln, Trp, Val, Phe, Lys, Ile
Helix breakers	Pro, Gly, Tyr, Asn
Indifferent residues	Asp, Thr, Ser, Arg, Cys

The β-Sheet Structures

The second major regular, repeating structure, the β structure, differs from the α helix in that the polypeptide chains are almost completely extended, as in Fig. 4-7(*a*), and hydrogen bonding occurs *between* polypeptide strands, rather than *within* a single strand, as shown in Fig. 4-7(*c*).

Adjacent chains can be aligned in the same direction (i.e., N terminal to C terminal) as in the *parallel β sheets*, or alternate chains may be aligned in opposite orientations as in the *antiparallel β sheets*, shown in Fig. 4-7(*c*). These structures often form extensive sheets, as shown in Fig. 4-7(*b*). Sometimes it is possible for several sheets to be stacked upon one another. Because the side chains tend to protrude above and below the sheet in alternating sequence, as shown in Fig. 4-7(*b*), the

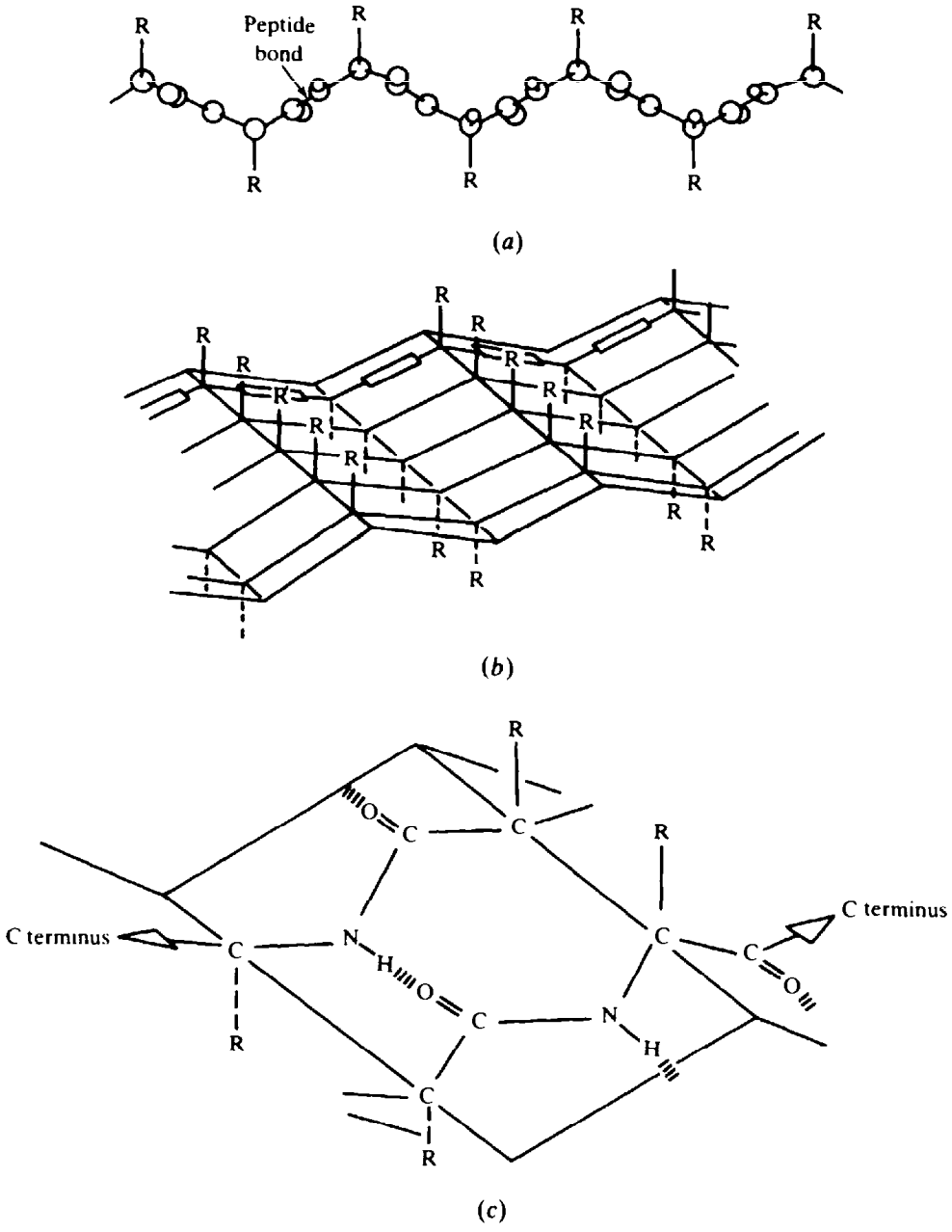


Fig. 4-7 β-Sheet structures: (*a*) polypeptide segment in an extended conformation; (*b*) sheet formed by the assembly of extended polypeptide chains side by side; (*c*) detail showing H bonding between adjacent polypeptide chains in an antiparallel β sheet.

β-sheet structures are favored by amino acids with relatively small side chains, such as alanine and glycine. Large, bulky side chains can lead to steric interference between the various parts of the protein chain.

EXAMPLE 4.13

Predict which regular, repeating structure is more likely for the two polypeptides (*a*) poly(Gly-Ala-Gly-Thr); (*b*) poly(Glu-Ala-Leu-His).

Polypeptide (*a*) is comprised largely of amino acid residues with small side chains. Except for Ala, none of the residues favors helix formation, and Gly destabilizes the α helix. Thus, this polypeptide is more likely to form β structures.

Polypeptide (*b*) is comprised of amino acid residues with bulky side chains that would destabilize β structures. However, all the amino acids are helix-stabilizing, and thus polypeptide (*b*) is more likely to form an α helix.

EXAMPLE 4.14

The structures of many globular proteins are comprised of elements of α helix, β structures, and disordered regions. It is instructive to represent the structures of these proteins using the stylized form shown in Fig. 4-8, in which α helices are represented by coiled ribbons and β structures are represented by arrows pointing in the N$\rightarrow$C direction. Parallel β sheets have their arrows pointing in the same direction; antiparallel β sheets have their arrows alternating. β proteins (e.g., retinol binding protein and antigen binding fragment) contain predominantly β sheet secondary structure, while α proteins (e.g., myoglobin) are largely composed of α helices. α/β proteins (e.g., triosephosphate isomerase) contain a mixture of α helices and β sheets.

The Collagen Triple Helix

The protein *collagen*, from skin and tendons, is composed of approximately 30 percent proline and hydroxyproline and 30 percent glycine. This protein has an unusual structure in which *three* chains, each with a conformation very similar to that of polyproline, are twisted about each other to make a triple helix. The three strands are hydrogen bonded to each other, through hydrogen bonds between the $-$NH of glycine residues and the $-$C=O groups of the other amino acids.

Question: Why is proline rarely found within α-helical segments?

The α-amino group of proline is a *secondary* amino group. When proline participates in peptide bonds through its amino group, there is no longer an amide hydrogen to participate in the hydrogen-bond stabilization of the α helix. In addition, because the side chain of proline is attached to the α-amino group, there is no free rotation about the N$-$C$_\alpha$ bond and proline cannot take up the correct conformation for an α-helical residue.

Although proline cannot participate in α-helical conformations, polypeptides composed only of proline can adopt a different type of helical conformation. This *polyproline helix* is *not* stabilized by hydrogen bonding, but rather by the steric mutual repulsion effects of the prolyl side chains. The polyproline helix is more extended than the α helix, with adjacent residues separated along the axis by 0.31 nm.

EXAMPLE 4.15

Why does collagen, with its polypeptide sequence that is largely (Gly-Pro-x-)$_n$, where x is another type of amino acid residue, form a triple helix, while polyproline does not?

In the collagen triple helix, every third residue is positioned toward the center of the helix and comes into close contact with another chain. Only glycine, with its single hydrogen atom side chain, is small enough to fit into this crowded space.

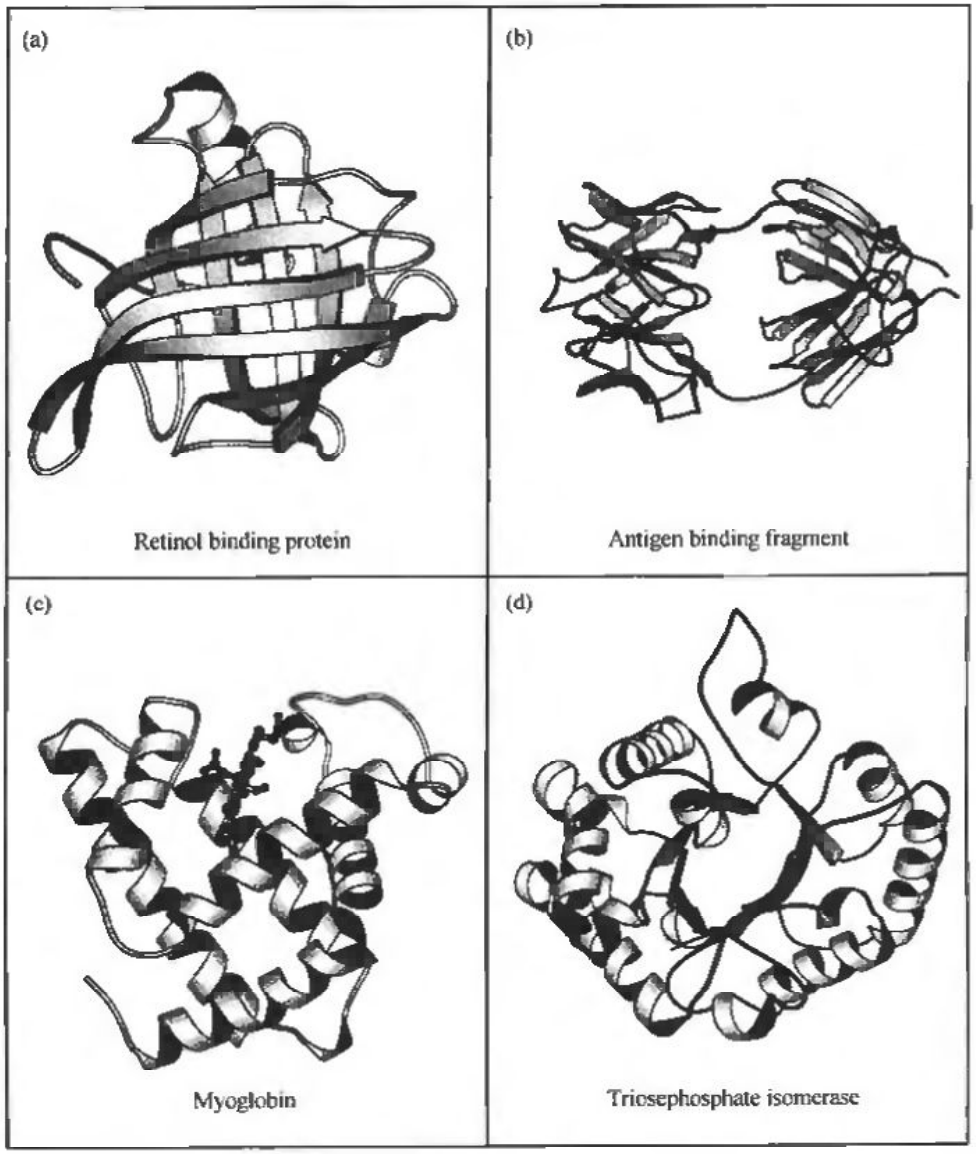


Fig. 4-8 Stylized representations of protein structures in which α helices are represented as coiled ribbons and β strands are represented by arrows pointing in the N→C direction. β proteins contain predominantly β-sheet structure (e.g., retinol binding protein and the antigen binding fragment of antibodies) while α proteins contain predominantly α helices (e.g., myoglobin). α/β proteins contain a mixture of α helix and β sheet (e.g., triosephosphate isomerase).

Optical Activity

Asymmetric molecules such as carbohydrates, amino acids, and proteins rotate the plane of polarization of plane-polarized light. The amount of rotation depends on the concentration of the substance and the path length of light in the sample (Chap. 2), in much the same way as for optical absorbances (Chap. 3). The amount of rotation (and, in fact, even the *direction* of rotation) also depends on the wavelength of the light. The dependence of the *specific rotation* $[\alpha]$ (the measured rotation per unit concentration and path length) on the wavelength of light is known as the *optical rotatory dispersion* (ORD).

The conformation of a protein introduces an additional source of asymmetry that affects the ORD spectrum. Helical regions in soluble proteins give rise to a particular ORD spectrum (see Fig. 4-9)

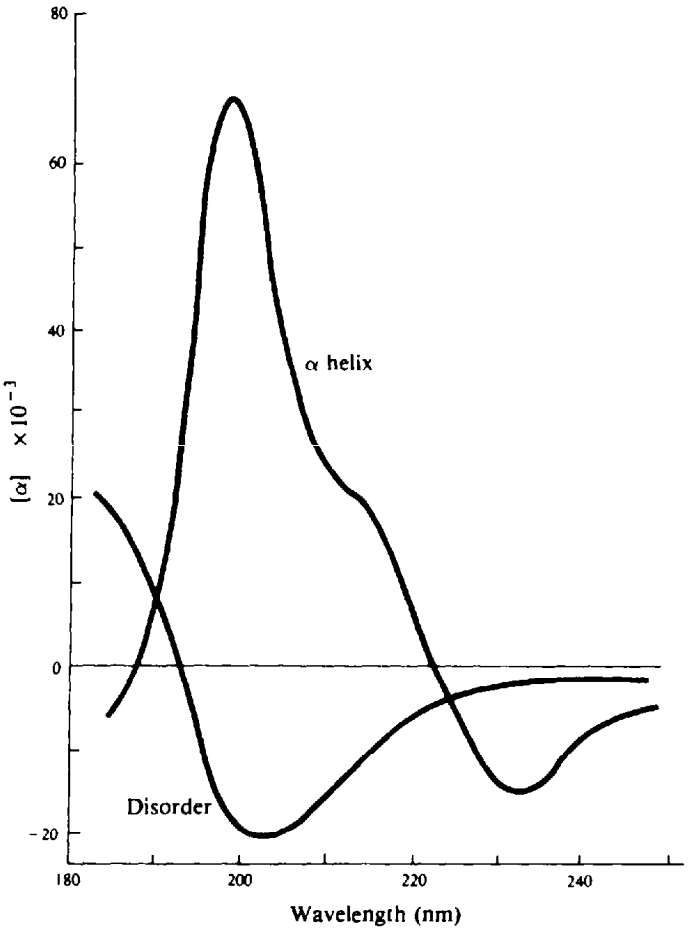


Fig. 4-9 ORD spectra of poly-D-lysine showing the spectra of α-helical and disordered conformations. (Units of $[\alpha]$ are degrees mL $dm^{-1} g^{-1}$.)

quite distinct from that of disordered regions. By comparing the ORD spectrum of an unknown protein with suitable standards of known conformation, it is possible to calculate the approximate proportions of the different types of structure in the protein.

Hierarchy of Protein Structure

It is possible to consider the structure of a protein on several levels as originally proposed by the Danish protein chemist, Kai Linderstrøm-Lang:

(*a*) *Primary structure*: the sequence of amino acids.

(*b*) *Secondary structure*: the regular, repeating folding pattern (such as the α helix and β pleated-sheet structures), stabilized by hydrogen bonds between peptide groups close together in the sequence. The secondary structure elements in a protein are often summarized in the form of a topology diagram (see Fig. 4-10) which defines in two dimensions their extent and relative orientations. These diagrams are often used to show relationships within families of proteins.

(*c*) *Supersecondary structure*: common repeating patterns of secondary structure that occur in many proteins. One common, recurring pattern is the β-α-β motif, which has a segment of β sheet, an intervening α helix, and a second segment of β-sheet hydrogen bonded to

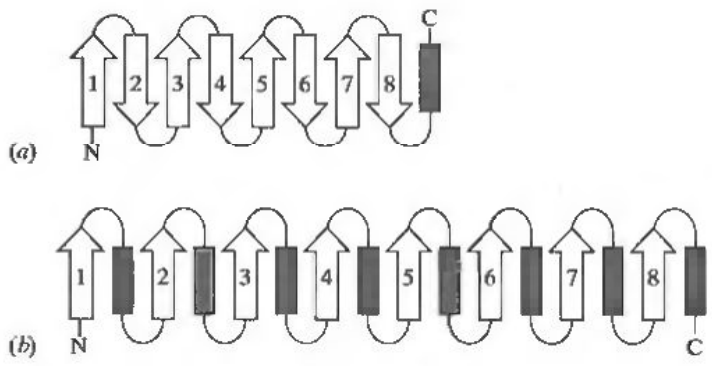


Fig. 4-10 Topology diagram for (*a*) retinol binding protein (RBP) and (*b*) triosephosphate isomerase (TPI). The arrows represent β strands (numbered from N to C) and the dark boxes represent α helices. Note from Fig. 4-8 that both of these proteins form a barrel structure comprised of eight β strands with the first strand hydrogen bonded to last strand in order to "close" the barrel. However, whereas the β strands are antiparallel in RBP, they are arranged in parallel in TPI and are surrounded by an outer layer of α helices which connect each β strand to the next in the barrel.

the first (Fig. 4-11). Other motifs include: the β hairpin, which comprises two antiparallel β strands connected by a tight reverse turn; the $\alpha\alpha$ motif, comprised of two closely packed antiparallel α helices; and β barrels, in which extended β sheets wrap around to form a continuous cylinder.

(*d*) *Tertiary structure*: for a globular protein, the tertiary structure can be thought of as the way that segments of secondary structure fold together in three dimensions, stabilized by interactions often far apart in the sequence. For those proteins with little or no detectable α helix or β structure, the tertiary structure can be considered as the way the protein folds in three dimensions, stabilized by interactions between distant parts of the sequence.

(*e*) *Domain structure*: domains are a common feature of many globular proteins, particularly where the molecular mass is more than 20 kDa. Larger proteins are often folded so that each domain is ~17 kDa. For example, the enzyme glyceraldehyde 3-phosphate dehydrogenase is folded into two domains, such that each has a separate function: one domain of ~16 kDa binds the cofactor NAD^+, while the other, catalytic domain, of ~21 kDa binds the substrate glyceraldehyde 3-phosphate (Chap. 11).

(*f*) *Quaternary structure*: the interaction between different polypeptide chains to produce an *oligomeric* structure, stabilized by noncovalent bonds only.

Families of Protein Structure

Now that a large number of globular protein structures have been determined experimentally, it is possible to group them, or each of their domains, into classes which have similar organization. The α/β family (which includes all of the glycolytic enzymes) contains repeating β-α-β motifs, whereas the $\alpha + \beta$ family contains unrelated β sheets and α helices. The α and β families contain almost exclusively α helix and β sheet, respectively. Most protein toxins belong to the family of small S-S bridged proteins, in which a small hydrophobic core is maintained by a relatively large number of disulfide bonds; for example, insect and scorpion defensins have only 38–40 residues but they each contain three disulfide bonds.

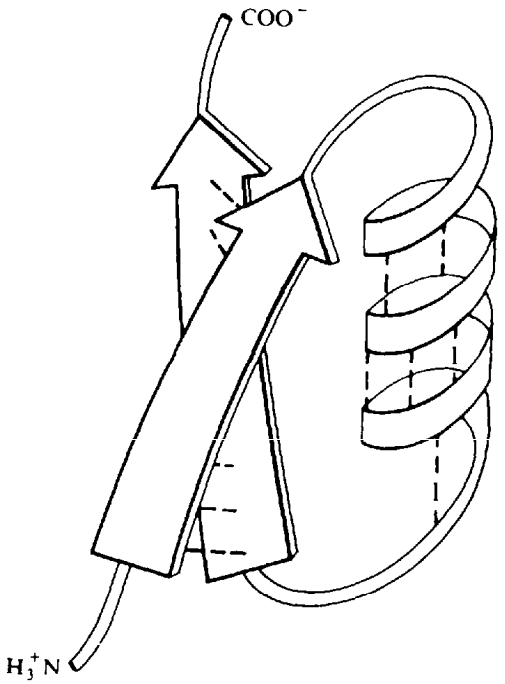


Fig. 4-11 Diagrammatic representation of the supersecondary β-α-β folding unit of a protein. β regions are represented by the arrows, while the α-helical segment is indicated by the coiled structure. Approximate positions of hydrogen bonds are shown with dashed lines.

EXAMPLE 4.16

For a spherical protein, how do (*a*) volume and (*b*) surface area depend on molecular weight?

(*a*) Given a uniform density, the volume V would be directly proportional to the mass, and therefore to the molecular weight; i.e., $V \propto M_r$.

(*b*) The surface area of a sphere is proportional to the square of the radius, while the volume is proportional to the cube of the radius. Thus, $A \propto V^{2/3}$. We have just seen that $V \propto M_r$; therefore, $A \propto M_r^{2/3}$.

The above calculation shows that the volume of a spherical protein will increase more rapidly with molecular weight than will the surface area. Thus, in order to accommodate all charged groups on the surface and all nonpolar groups in the interior, three strategies are possible:

1. Larger proteins may show altered composition, with increasing proportions of nonpolar amino acid side chains occupying the increased interior volume.
2. Larger proteins may fold into separate *domains*, each domain being a *globular folding unit* with its own interior and surface, and with an interconnecting strand of backbone linking the domains.
3. Large proteins may fold into more elongated or rodlike shapes.

4.5 SEQUENCE HOMOLOGY AND PROTEIN EVOLUTION

The protein *myoglobin* serves as an oxygen binder in muscle. It was among the first to be studied with the aid of x-ray crystallography, which revealed a compact globular structure comprising eight segments of α helix linked via short nonhelical segments.

Hemoglobin, the oxygen-carrying protein of vertebrate blood, is similar in structure to myoglobin. However, hemoglobin is composed of four chains; i.e., it has a *quaternary structure*. The four chains

are held together in a particular geometrical arrangement by noncovalent interactions. There are two major types of hemoglobin polypeptide chains in normal adult hemoglobin, the α chain and the β chain.

The polypeptide chain of myoglobin and the two chains of hemoglobin are remarkably similar, both in primary and in tertiary structure. The two proteins are said to be *homologous*. Myoglobin has 153 residues, the hemoglobin α chain has 141, and the hemoglobin β chain 146. The sequence is identical for 24 out of 141 positions for the human proteins, and many of the differences show *conservative replacement*. This means that an amino acid residue in one chain has been replaced by a chemically similar residue at the corresponding position in another chain, for example, the replacement of glutamate by aspartate.

Comparison of the amino acid sequences of hemoglobin and myoglobin chains from different species of animals shows that the chains from related species are similar. The number of differences increases with phylogenetically more separated species. On the assumption that proteins evolve at a constant rate, the number of differences between two homologous proteins will be proportional to the time of divergence in evolution of the species.

EXAMPLE 4.17

Draw an evolutionary tree for the human, rabbit, silkworm, and *Neurospora* (fungus) by using the data in the following table for differences in the respective cytochrome *c* sequence.

	Number of Sequence Differences			
	Human	Rabbit	Silkworm	*Neurospora*
Human	0	11	36	71
Rabbit		0	35	70
Silkworm			0	69
Neurospora				0

These data allow us to construct an evolutionary tree, with branch lengths approximately proportional to the number of differences between the species (see Fig. 4-10). The human and rabbit show most similarity and therefore are connected by short branches. The silkworm cytochrome *c* is closer to both mammalian forms than it is to *Neurospora*, and so should be connected to the mammalian junction. The length of the silkworm branch will be approximately three times the length of the rabbit and human branches. Finally, *Neurospora* shows approximately the same number of differences with all the animal species, and therefore can be joined to their common branch.

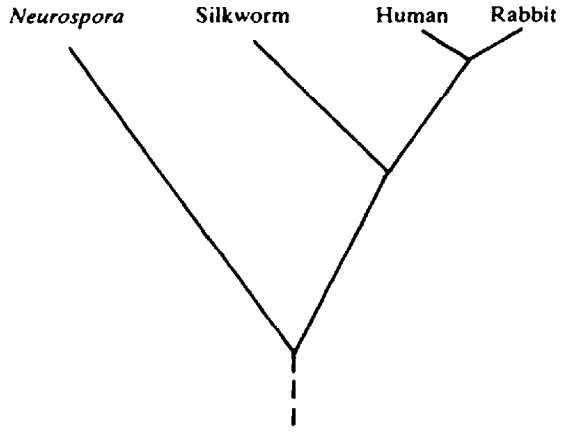


Fig. 4-12 Phylogenetic tree constructed from amino acid sequences of cytochrome *c*.

4.6 METHODS FOR PROTEIN STRUCTURE DETERMINATION

Most of our structural information comes from x-ray crystallographic analysis of protein crystals and from the use of nuclear magnetic resonance spectroscopy in solution. Each of these techniques has advantages and limitations which makes them suitable for a complementary range of problems. The first protein structure determined at a sufficient resolution to trace the path of the polypeptide chain was that of myoglobin in 1960. Since that time many thousands of structures corresponding to hundreds of different proteins have been determined. The coordinates of the atoms in many protein and nucleic acid structures are available from the Protein Data Bank, which may be accessed via the *Internet* or *World Wide Web* (http://www.pdb.bnl.gov).

Question: What generalizations or folding rules can be drawn from the known data base of experimentally determined protein structures?

1. Most electrically charged groups are on the surface of the molecule, interacting with water. Exceptions to this rule are often catalytically important residues in enzymes, which may well be partially stabilized by specific polar interactions within a hydrophobic portion of the molecule.
2. Most nonpolar (e.g., hydrocarbon) groups are in the interior of the molecule, thus avoiding thermodynamically unfavorable contact with water. Exceptions to this may function as specific binding sites on the surface of the molecule for other proteins or ligands.
3. Maximal hydrogen bonding occurs within the molecule.
4. Proline often *terminates* α-helical segments.

---

x-Ray Crystallography

x-Rays have a wavelength which approximates the distance between bonded atoms in a molecule (e.g., a carbon-carbon single bond is 1.5 Å and the wavelength of x-rays from a tube with a copper target is 1.54 Å). When a beam of monochromatic x-rays interacts with a molecule it is diffracted and the diffraction pattern contains information about the location of the atoms in the diffracting molecule. If the sample is in solution the individual molecules of the protein will be randomly orienting and only very low resolution information on the overall shape of the molecules will be obtained. Fibrous proteins and nucleic acids can be induced to form oriented gels or fibers. x-Ray diffraction from such samples was used to define regular repeating spacings which characterized the α helix and β sheet in proteins and the A and B forms of DNA and RNA. In these cases x-ray diffraction is capable of being used only to refine a proposed geometric model but *not* to uniquely and objectively define the structure.

The most powerful use of x-ray diffraction is in conjunction with single crystals of the sample. Structures of molecules ranging in size from a few atoms to viruses with tens of thousands of atoms have been solved. In a single crystal all the molecules lie on a three-dimensional lattice with fixed relative positions and orientations. The intensities of each point in the diffraction pattern can be measured and therefore information regarding the full three-dimensional structure can be obtained.

Question: What are the advantages and limitations of protein crystallography?

1. The technique is applicable to a wide range of samples which includes proteins, enzymes, nucleic acids and viruses, and is not limited by the size of the molecules.
2. Not all proteins can be crystallized. This is especially true of intrinsic membrane proteins which need to be solubilized in the presence of detergents.

3. Most protein crystals do not diffract to true atomic resolution. Due to an inherent lack of perfection in protein crystals the high angle component of the diffraction pattern, which contains the fine detail, is lost. For this reason the protein sequence must generally be known.

Nuclear Magnetic Resonance Spectroscopy

NMR spectroscopy has only developed as a technique for protein structure determination in the past decade; the first complete three-dimensional protein structure solved using this technique was presented in 1986. The NMR phenomenon derives from the fact that the energy levels of nuclei with nonzero spin become unequal when the nuclei are placed in a magnetic field. The energy of the nuclei can therefore be perturbed (i.e., moved between energy levels) by the application of radiofrequency pulses whose wavelength corresponds to the gap between these energy levels. The ability to perturb the nuclei in this way enables the observation of scalar and dipolar nuclear interactions. The first type of interaction allows information to be obtained about protein torsion angles, while dipolar interactions reveal information about interproton distances. The structure of the protein is calculated using the amino acid sequence of the protein in combination with these measurements of dihedral angles and interproton distances.

It should be noted that NMR studies are performed on proteins in their native solution state. Since the protein molecules are moving around in solution, the NMR technique measures scalar quantities (torsion angles and interproton distances); this is fundamentally different from x-ray crystallography in which the static crystal lattice allows a vector "image" of the molecule to be obtained. Since NMR studies are performed on proteins in solution, the technique can be used to probe intricate details of the dynamics of the protein.

Question: What are the advantages and limitations of protein NMR spectroscopy?

1. The technique is applicable to both proteins and nucleic acids in their native *solution* state.
2. There is an upper size-limit of 35–40 kDa. However, this means that most protein domains can be studied using NMR spectroscopy.
3. The final resolution is not as good as x-ray crystallography but much useful information can be obtained about the dynamics of the protein.

EXAMPLE 4.18

What is the relevance of a structure determined in a crystalline state to its native structure in solution?

Protein crystals have a solvent content of 40–70% by volume. Thus, there are few direct intermolecular contacts between molecules which affect the structure. Structures of the same protein determined in different crystal forms, and therefore containing different packing contacts, have generally been found to be the same. This is also true of structures determined using both NMR spectroscopy and protein x-ray crystallography. It has also been found that some enzymes are fully active in the crystalline state, implying that crystallization has not altered the native conformation.

EXAMPLE 4.19

If one had a 20 kDa protein, what preliminary steps could be taken to decide whether it was suitable for a structural study using protein crystallography and/or NMR spectroscopy?

(*a*) Gel filtration or analytical ultracentrifugation (sedimentation equilibrium) experiments could be performed to determine if the protein self-associated to form oligomers. Oligomerization of the protein

would preclude structure determination using NMR spectroscopy, as even a dimer would be at the very upper molecular-weight limit of this technique.

(*b*) If the protein does not oligomerize, NMR studies could begin so long as a 250 μL sample of ~1 m*M* concentration or greater could be obtained.

(*c*) Even if oligomerization occurred, crystallization studies could proceed so long as the oligomers were all identical. Methods that explore large numbers of different parameters (such as buffer and pH) can be used to find suitable crystallization conditions. Since proteins are least soluble at their isoelectric point (pH_I), this is often a good pH at which to begin crystallization trials.

Solved Problems

PURIFICATION AND CHARACTERIZATION OF PROTEINS

4.1. A pure heme protein was found to contain 0.426 percent of iron by weight. What is its minimum molecular weight?

SOLUTION

The minimum molecular weight is the molecular weight of a molecule containing only a single iron atom. Thus, if 0.426 g of iron is contained in every 100 g of protein, then 1 mol (56 g) iron is contained in

$$\frac{100 \times 56}{0.426} = 13{,}145 \text{ g of protein}$$

The mass of protein containing one gram atom of iron is 13,145 g, which therefore represents the molar mass of the protein. The minimum molecular weight is thus 13,145. This is close to the molecular weight of the heme protein cytochrome *c*, and is thus a reasonable value. However, had the molecule of protein contained more than one atom of iron, the molecular weight would have been a multiple of 13,145.

4.2. Threonine constitutes 1.8 percent by weight of the amino acid content of insulin. Given that the molecular weight of threonine is 119, what is the minimum molecular weight of insulin?

SOLUTION

Since water is lost during the condensation of amino acids to form peptide bonds, the *residue weight* of threonine will be 119 − 18 = 101. If we assume a single threonine residue per molecule, then 101 represents 1.8 percent of the molecular weight, and

$$M_r = \frac{101}{0.018} = 5{,}600$$

Physical measurements in dissociating solvents confirm this value.

4.3. Hemoglobin A (the major, normal form in humans) has an isoelectric point of pH 6.9. The variant hemoglobin M has a glutamate residue in place of the normal valine at position 67 of the α chain. What effect will this substitution have on the electrophoretic behavior of the protein at pH 7.5?

SOLUTION

Since pH 7.5 is above the isoelectric point of hemoglobin A, the protein carries a negative charge and will migrate to the anode. At pH 7.5 the glutamate side chain has a negative charge, while valine

is uncharged. Hemoglobin M, therefore, carries an additional negative charge at pH 7.5 and will migrate *faster* toward the anode.

4.4. The proteins ovalbumin ($pH_I = 4.6$), urease ($pH_I = 5.0$), and myoglobin ($pH_I = 7.0$) were applied to a column of DEAE-cellulose at pH 6.5. The column was eluted with a dilute pH 6.5 buffer, and then with the same buffer containing increasing concentrations of sodium chloride. In what order will the proteins be eluted from the column?

SOLUTION

At pH 6.5, both ovalbumin and urease are negatively charged, and will bind to the DEAE-cellulose. Myoglobin has a positive charge at pH 6.5 and will be eluted immediately. As the salt concentration is raised, electrostatic interactions are weakened; urease will be eluted next, and ovalbumin will be eluted last.

4.5. An enzyme of $M_r = 24{,}000$ and $pH_I = 5.5$ is contaminated with a protein of similar molecular weight, but with $pH_I = 7.0$, and another protein of $M_r = 100{,}000$ and $pH_I = 5.4$. Suggest a purification strategy.

SOLUTION

Gel filtration will allow the high-molecular-weight contaminant to be removed. The remaining mixture of lower-molecular-weight proteins can be separated by ion-exchange chromatography, as described in Prob. 4.4.

4.6. (*a*) What peptides would be released from the following peptide by treatment with trypsin?

Ala-Ser-Thr-Lys-Gly-Arg-Ser-Gly

(*b*) If each of the products were treated with fluoro-2,4-dinitrobenzene (FDNB) and subjected to acid hydrolysis, what DNP-amino acids could be isolated?

SOLUTION

(*a*) Trypsin hydrolyses peptides at the carboxyl side of lysine and arginine residues. The resulting peptides would be Ala-Ser-Thr-Lys, Gly-Arg, and Ser-Gly.

(*b*) Treatment with FDNB and hydrolysis will liberate DNP derivatives of the N-terminal amino acids: DNP-Ala, DNP-Gly, and DNP-Ser. Note that the ε-amino group of lysine can also react with FDNB; however, the ε-DNP derivative of lysine can be distinguished from the α-DNP derivative by its chromatographic behavior.

4.7. The following data were obtained from partial cleavage and analysis of an octapeptide:

Composition:	Ala, Gly_2, Lys, Met, Ser, Thr, Tyr
CNBr:	(1) Ala, Gly, Lys, Thr
	(2) Gly, Met, Ser, Tyr
Trypsin:	(1) Ala, Gly
	(2) Gly, Lys, Met, Ser, Thr, Tyr
Chymotrypsin:	(1) Gly, Tyr
	(2) Ala, Gly, Lys, Met, Ser, Thr
N terminus:	Gly
C terminus:	Gly

Determine the sequence of the peptide.

SOLUTION

A set of overlapping peptides can be prepared from the above data by making use of the fact that one CNBr peptide must end in methionine, one tryptic peptide must end in either lysine or arginine, and one chymotryptic peptide must end in an aromatic amino acid. The sequence of the peptide is, therefore, Gly-Tyr-Ser-Met-Thr-Lys-Ala-Gly.

4.8. The enzyme *carboxypeptidase A* hydrolyzes amino acids from the C-terminal end of peptides, provided that the C-terminal residue is not proline, lysine, or arginine. Fig. 4-13 shows the sequential release of amino acids from a protein by means of treatment with carboxypeptidase A. Deduce the sequence at the C terminus.

SOLUTION

The sequence at the C terminus is:

-Ser-His-Ile

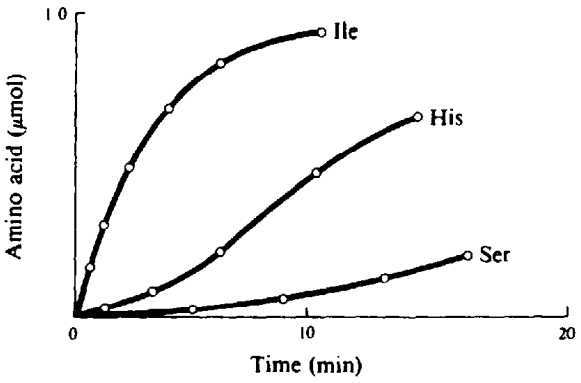


Fig. 4-13

4.9. The artificial polypeptide poly-L-glutamate forms α helices in solution at pH 2, but not at pH 7. Suggest an explanation.

SOLUTION

At pH 2, the side chains of poly-L-glutamate are largely uncharged and protonated. However, at pH 7 they are negatively charged. The negative charges lead to mutual repulsion and destabilization of the α helix.

4.10. Poly-L-glutamate in 100 percent α-helical form shows a trough in the ORD spectrum with $[\alpha]_{233} = -15{,}000°$; in the random-coil form, $[\alpha]_{233} = -1{,}000°$ (the subscript 233 refers to the wavelength of light used). Calculate the proportion of α helix in a protein for which $[\alpha]_{233} = -7{,}160°$, assuming the presence of α-helical and disordered regions only.

SOLUTION

100 percent α helix corresponds to $[\alpha]_{233} = -15{,}000°$; 0 percent α helix corresponds to $[\alpha]_{233} = -1{,}000°$. Therefore

$$\% \text{ helix} = \frac{-[\alpha]_{233} - 1{,}000}{15{,}000 - 1{,}000} \times 100\% = \frac{6{,}160}{14{,}000} \times 100\% = 44\%$$

4.11. Myoglobin in solution at pH values above 6.0 has a value of specific rotation $[\alpha]_{233} = -12{,}000°$. However, on adjusting the solution to pH 2.0, the specific rotation changes to $-2{,}000°$. Explain the change.

SOLUTION

Using the method of Prob. 4.10 the helical content of myoglobin can be calculated at both pH values, if we assume the presence of α-helical and disordered regions only. At pH 6.0, the helical content is 79 percent. However, at pH 2.0 the helical content is only 7 percent. We may conclude, then, that at low pH, the protein has lost its α-helical conformation; i.e., it has been denatured.

4.12. Partial sequence determination of a peptide gave the following:

-Gly-Pro-Ser-Gly-Pro-Arg-Gly-Leu-Hyp-Gly-

What conclusions can be made about the possible conformation of the protein from which this peptide was derived?

SOLUTION

This sequence closely resembles that of collagen. In particular, the repeating (Gly-Pro-X) pattern and the occurrence of hydroxyproline are characteristic of collagen. It is possible, then, that the peptide was derived from a protein resembling collagen.

4.13. Insulin possesses two polypeptide chains, A and B, linked by disulfide bonds. Upon denaturation and reduction of insulin, followed by reoxidation, only 7% recovery of activity was obtained. This is the level of activity expected for random pairing of disulfide bridges. How can these data be reconciled with the hypothesis that the amino acid sequence directs protein folding?

SOLUTION

Insulin is synthesized as *proinsulin*. After synthesis and folding, a section of the molecule (the C peptide) is excised, leaving the A and B peptides connected via disulfide bridges. Thus, native insulin, lacking the C peptide, lacks some of the information necessary to direct the folding process.

4.14. Insulin and hemoglobin are both proteins that comprise more than one polypeptide chain. Contrast the interactions between the component polypeptide chains of the two proteins.

SOLUTION

The two chains of insulin, fragments of what was originally a single chain, are held together by covalent disulfide bonds. Hemoglobin has four polypeptide chains, held together by noncovalent interactions only.

4.15. There are four disulfide bonds in ribonuclease. If these are reduced to their component sulfhydryl groups and allowed to reoxidize, how many different combinations of disulfide bonds are possible?

SOLUTION

In forming the first disulfide bond, a cysteine residue may pair with any one of the remaining seven. In forming the second bond, a cysteine residue may pair with one of the remaining five, and for the third bond, a cysteine residue may pair with one of three. Once three bonds have formed, there is only one way to form the last bond. Consequently, the number of possibilities is

$$7 \times 5 \times 3 = 105$$

Thus, the likelihood of forming the correct disulfide bonds by chance alone is

$$\frac{1}{105} = 0.0095, \text{ or } 0.95\%$$

4.16. The glycine residue at position 8 in the sequence of insulin has torsion angles $\phi = 82°$, $\psi = -105°$, which lie in the unfavorable region (marked X) of the Ramachandran plot in Fig. 4-5. How is this possible?

SOLUTION

Glycine has a very small side chain, a single hydrogen atom. The plot of Fig. 4-5 was determined for alanine, which has a methyl group as the side chain. Thus, there are conformations allowed for glycine that are not possible for the methyl group of alanine.

4.17. Using the data of Example 4.7, estimate the molecular weight of an enzyme for which the elution volume was 155 mL.

SOLUTION

From Fig. 4-2, the value of log M_r corresponding to 155 mL elution volume is 4.8. The molecular weight is therefore 63,000, providing that the density and shape of the enzyme are similar to those of the calibration standards.

4.18. An enzyme examined by means of gel filtration in aqueous buffer at pH 7.0 had an apparent molecular weight of 160,000. When examined by gel electrophoresis in SDS solution, a single band of apparent molecular weight 40,000 was formed. Explain these findings.

SOLUTION

The detergent SDS causes the dissociation of quaternary structures and allows the determination of molecular weight of the component subunits. The data suggest that the enzyme comprises four identical subunits of $M_r = 40{,}000$, yielding a tetramer of $M_r = 160{,}000$.

4.19. During an attempt to determine the molecular weight of the milk protein β-lactoglobulin, by means of gel filtration, the following data were obtained with different sample concentrations:

Protein Concentration	Apparent M_r
$10\ g\ L^{-1}$	36,000
$5\ g\ L^{-1}$	35,000
$1\ g\ L^{-1}$	32,000
$0.1\ g\ L^{-1}$	25,000

Electrophoresis in polyacrylamide gels containing SDS led to an apparent molecular weight of 18,000, consistent with the known amino acid sequence of this protein. Explain these data.

SOLUTION

These data show that the polypeptide chain of β-lactoglobulin has a molecular weight of 18,000, and that in high concentration, the protein exists as a dimer of $M_r = 36{,}000$. However, on dilution, the dimer, maintained by reversible, noncovalent interactions, undergoes a partial dissociation: $A_2 \rightleftharpoons 2A$.

4.20. Calculate M_r of a protein, given the following experimental data obtained at 20°C.

$$s = 4.2 \times 10^{-13}\ s$$
$$D = 1.2 \times 10^{-10}\ m^2\ s^{-1}$$
$$\bar{v} = 0.72\ mL\ g^{-1}$$
$$\rho = 0.998\ g\ mL^{-1}$$

SOLUTION

From Eq. (*4.1*)

$$M = \frac{RTs}{D(1 - \bar{v}\rho)}$$

where R (gas constant) $= 8.314\,\text{J K}^{-1}\,\text{mol}^{-1}$
T (absolute temperature) $= 293\,\text{K}$

$$M = \frac{8.314\,\text{J K}^{-1}\,\text{mol}^{-1} \times 293\,\text{K} \times 4.2 \times 10^{-13}\,\text{s}}{1.2 \times 10^{-10}\,\text{m}^2\,\text{s}^{-1} \times (1 - 0.72 \times 0.998)}$$

$$= \frac{10{,}231 \times 10^{-13}\,\text{kg m}^2\,\text{s}^{-2}\,\text{mol}^{-1}\,\text{s}}{0.338 \times 10^{-10}\,\text{m}^2\,\text{s}^{-1}}$$

$$= 30\,\text{kg mol}^{-1} \text{ (to two significant figures)} \quad \text{or} \quad 30{,}000\,\text{g mol}^{-1}$$

The molecular weight is thus 30,000.

Supplementary Problems

4.21. List (*a*) the major proteins of muscle, (*b*) the major protein of skin, connective tissue, and bone, and (*c*) the major protein of hair and feathers.

4.22. A pure heme protein was found to contain 0.326 percent iron. If the molecule contains only one iron atom, what is its molecular weight?

4.23. In what direction (toward the anode, toward the cathode, or toward neither) will the following proteins move in an electric field?

(*a*) Serum albumin (I.P. = 4.9) at pH 8.0

(*b*) Urease (I.P. = 5.0) at pH 3.0; pH 9.0

(*c*) Ribonuclease (I.P. = 9.5) at pH 4.5; pH 9.5; pH 11.0

(*d*) Pepsin (I.P. = 1.0) at pH 3.5; pH 7.0; pH 9.5

4.24. In what order would the following globular proteins emerge on gel filtration of a mixture on Sephadex G-200: ribonuclease (M_r = 12,000); aldolase (M_r = 159,000); hemoglobin (M_r = 64,000); β-lactoglobulin (M_r = 36,000); and serum albumin (M_r = 65,000)?

4.25. Distinguish between the terms *primary*, *secondary*, and *tertiary* structures.

4.26. Of the following amino acid residues—methionine, histidine, arginine, phenylalanine, valine, glutamine, glutamic acid—which would you expect to find on the (*a*) surface of a protein and which would you expect to find (*b*) in the interior?

4.27. What functions would you expect to be served by residues such as (*a*) phenylalanine at the protein surface or (*b*) aspartic acid in the interior?

4.28. What is meant by a *domain* of protein structure? Give an example of a domain in a real protein.

4.29. What is meant by the statement that a particular conformation of an amino acid residue lies in an *unfavorable region* of the Ramachandran plot?

4.30. (*a*) Why does urea cause protein denaturation? (*b*) Why do our kidneys not denature in the presence of urinary urea?

4.31. (*a*) What are the important noncovalent interactions within proteins? (*b*) How do weak interactions result in a stable structure?

4.32. The *pitch* (p) of a helix is defined as $p = dn$, in which n is the number of repeating units per turn and d is the distance along the helix axis per repeating unit. Therefore, the pitch is a measure of the distance from one point on the helix to the corresponding point on the next turn of the helix.

(*a*) What is the pitch of an α helix and the distance per residue?

(*b*) How long would myoglobin be if it were one continuous α helix?

(*c*) How long would myoglobin be if it were one strand of a β sheet?

(*d*) How long would myoglobin be if it were fully extended (distance/residue = 0.36 nm)?

4.33. Predict which of the following polyamino acids will form α helices and which will form no ordered structures in solution at room temperature.

(*a*) Polyleucine, pH = 7.0

(*b*) Polyisoleucine, pH = 7.0

(*c*) Polyarginine, pH = 7.0

(*d*) Polyarginine, pH = 13.0

(*e*) Polyglutamic acid, pH = 1.5

(*f*) Polythreonine, pH = 7.0

4.34. What forces hold protein subunits in a quaternary structure?

4.35. Poly-L-proline can form a single-strand helix that is similar to that of a single strand of the collagen triple helix, but it cannot form a triple helix. Why not?

4.36. Poly(Gly-Pro-Pro) is capable of forming triple helices. Why?

4.37. Compare and contrast the structures of (*a*) insulin, (*b*) hemoglobin, and (*c*) collagen, all of which are proteins consisting of two or more chains but held together by different types of bonds.

4.38. What are the reasons for the marked stability of an α helix?

4.39. (*a*) In what important ways do the α helix and β structure differ? (*b*) How are they similar?

Chapter 5

Proteins: Supramolecular Structure

5.1 INTRODUCTION

Although proteins are large molecules they are small compared with a cell and even with *supramolecular structures* which may be part of a cell, such as plasma and organelle membranes, ribosomes, chromosomes, filaments, enzyme complexes and viruses (Chap. 1). Supramolecular structures are also prominent outside cells and are, for example, essential components of connective tissues such as tendon, ligament, cartilage and bone. Supramolecular structures can consist of a variety of different types of molecule from the small (such as membrane lipids) to macromolecules (such as proteins, DNA and RNA).

These structures are involved in a huge variety of cellular and physiological processes and it is not possible for an organism, or any cell within an organism, to function properly without them. Some supramolecular structures, such as the *cytoskeletal networks* that support the microvilli of the brush border cells lining the small intestine, appear to be relatively static: i.e., their structure remains relatively unchanged throughout the lifetime of the cell. Most, however, are remarkably dynamic in several senses. (1) The turnover of components within them may be high (such as occurs for most membranes and some *extracellular matrices*). (2) They may perform mechanical work (such as cilia, flagella and muscle fibers). (3) They may respond to mechanical work or some other external stimulus (such as the oscillations of stereocilia on the surface of the inner hair cells in the ear in response to sound, and the remarkable deformations of human red cells in response to shear stress and hydraulic lines of flow as the cells navigate the circulation). (4) They may be built for a specific purpose and then dismantled (such as the needle-like *acrosomal process* of sea-urchin sperm which is built to assist the sperm in penetrating the extracellular material and membrane of the ovum, and is then destroyed).

This Chapter will highlight some of the features relating to the structure and construction of supramolecular structures that are primarily protein based. Examples of supramolecular structures found outside cells (extracellular matrices) and within cells (cytoskeletal networks) will be given that emphasize the relationships between the structure and function of these networks, the role of their frequently dynamic nature, and the genetic and congenital errors that can lead to, or be associated with, disease.

EXAMPLE 5.1

The dynamic features listed above are often combined in many types of supramolecular structure. The *mitotic* and *meiotic spindles* are excellent examples. The spindles are built for the purpose of separating the chromosomes to the daughter cells of a dividing cell and are then dismantled. The separation of chromosomes, of course, requires mechanical work and the expenditure of energy by components in the spindle.

5.2 ASSEMBLY OF SUPRAMOLECULAR STRUCTURES

Many supramolecular structures are formed largely by the stepwise noncovalent association of macromolecules, such as proteins. The processes of assembly are governed by the same chemical and physical principles that govern protein folding and the formation of quaternary structures (see Chap. 4). The driving force for the assembly process generally depends on the formation of a multitude of relatively weak hydrophobic, hydrogen and ionic bonds that occur between complementary sites on subunits which are in van der Waals contact with each other. In addition, covalent

crosslinking (such as disulfide bonding between cysteines on neighboring subunits) can also occur. The specificity of complementary binding sites means that the subunits in the aggregate bear a fixed orientation to each other.

Conformational changes within and between subunits are frequently essential to the assembly process. These conformational changes are often critical to the overall stability of the final structure with respect to the individual components. They can also strongly contribute to the highly cooperative nature of the assembly of supramolecular structures as well as contribute to the proper orientation of subunits with respect to one another.

In some cases, the assembly process is brought about by the association of a number of *identical subunits* to form a complex structure. This stepwise assembly process has certain advantages. First, it reduces the amount of genetic information needed to code for a complex structure, and second, it allows mistakes to be circumvented if faulty subunits can be excluded from the final structure (see Prob. 5.1).

When all of the information for assembly of a supramolecular structure is contained within the component molecules themselves, the process is termed *self-assembly*.

EXAMPLE 5.2

Ribosomes are large macromolecular complexes whose components contain all the information necessary for self-assembly. The *E. coli* ribosome has a sedimentation coefficient of 70 S and consists of two subunits (50 S and 30 S) with a total mass of 2.8×10^6 Da and with 58 different components. Three of these components are RNA molecules that together comprise 65 percent of the mass and they act as a framework or template for the ordering of the different proteins. When the pure dissociated components are mixed together in the proper order under the correct conditions they spontaneously reassemble to form a fully active ribosome (Fig. 5-1).

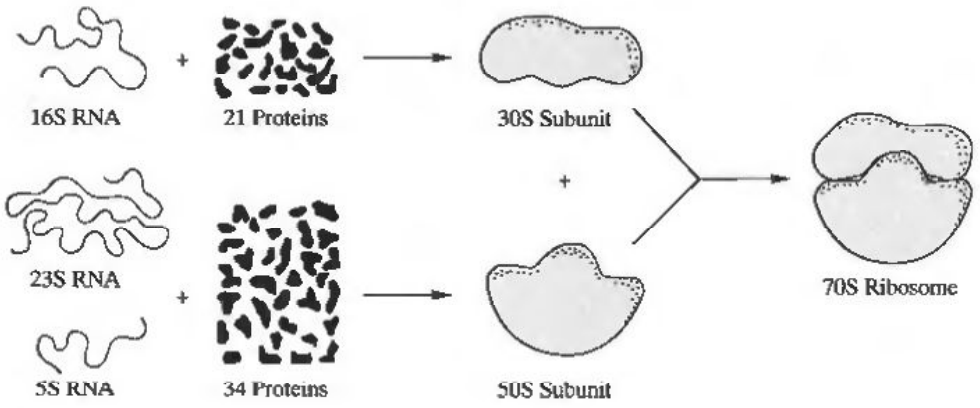


Fig. 5-1 Steps in the assembly of an *E. coli* ribosome.

Certain viruses, notably *tobacco mosaic virus* (TMV), can also self-assemble. A TMV particle can be dissociated into its component proteins and RNA and then reassembled into infective virus particles on mixing the components together again.

EXAMPLE 5.3

TMV consists of a cylindrical coat of 2,130 identical protein subunits enclosing a long RNA molecule of 6,400 nucleotides. In 1955, it was shown that the coat protein subunits and the RNA could be dissociated but would, under appropriate conditions, spontaneously self-assemble to reform fully active virus particles. This process is multistage, the critical intermediate being a 34-unit *two-layered* protein disc which, upon binding the RNA, is converted to a *helical* structure with 16.33 protein subunits per turn (Fig. 5-2). In the absence of the RNA, the protein may be polymerized into helical tubes of *indefinite* length. The presence of the RNA aids the polymerization process and results in a virus particle with a fixed length of 300 nm.

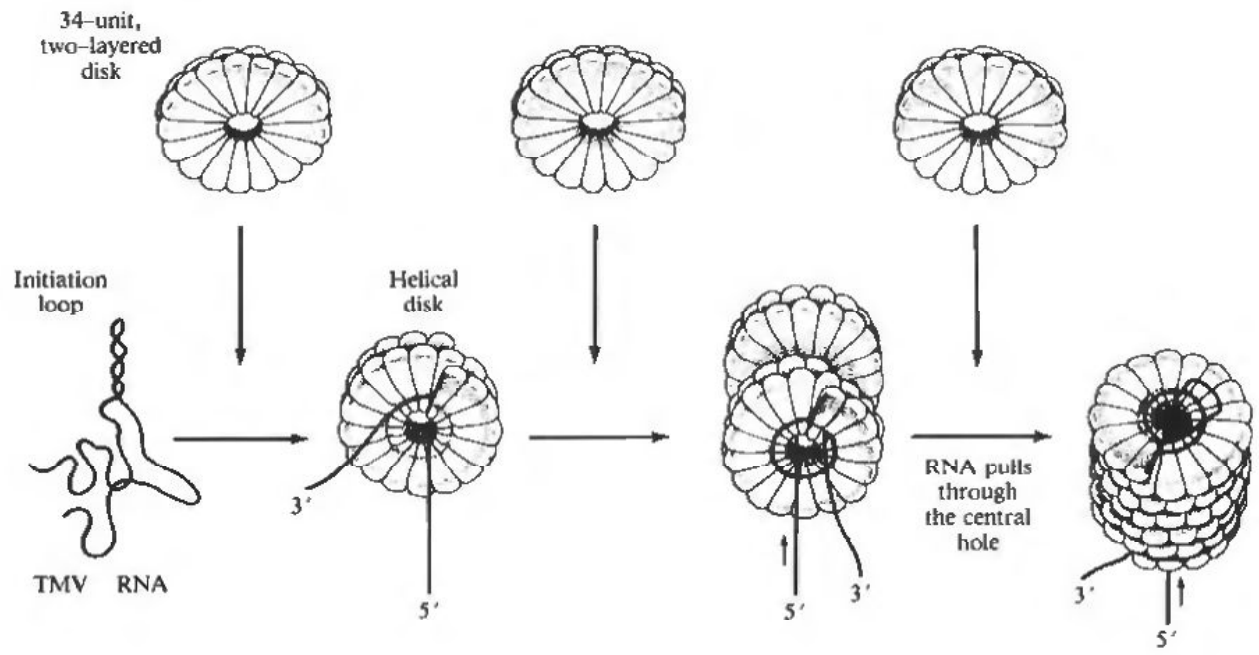


Fig. 5-2 The assembly of tobacco mosaic virus.

The construction of a viral coat from multiple copies of a comparatively small protein represents a huge saving in genetic information. Even when the virus is more complex with several coat proteins, the saving is still large. For instance, in the polio virus coat a 115 kDa precursor protein associates to form a pentamer. Each of the polypeptide chains of these pentamers is then cleaved by proteolysis to yield four proteins. Twelve of these proteolytically processed pentamers then associate to form the complete coat. Note that, in this case, self-assembling of components is not the only factor required for the full and proper formation of the final structure. In addition, input from *outside* the completed system, in the form of processing by proteases, is needed.

Other examples of external input include the use of a scaffold or template to order the assembly process, and the necessity for the information already stored in a structure to be used in the formation of a new copy. Examples of these include the assembly of the bacteriophage T4 virus, which involves the use of template proteins, and the formation of new mitochondria which arise from the growth and division of preexisting mitochondria.

EXAMPLE 5.4

The bacteriophage T4 is a complex virus capable of infecting certain bacteria. The virus' protein coat (head, tail and tail fibers) contains 40 structural proteins. T4 illustrates well the spatial and temporal control of the stepwise assembly process and the role of external input in regulating that process. A further 13 proteins are required for assembly *but do not appear in the completed virus particle*. Three of these act as a transient template to promote the formation of the tail baseplate (Fig. 5-3). Another one is a protease which cleaves the major protein of the head (from 55 kDa to 45 kDa) but only after the head has been partially assembled. It is only when the tail reaches its correct length that the cap protein is placed on top allowing the completed head to become attached (Fig. 5-3). Finally, the intact tail fibers are added at the baseplate, and this requires an enzyme-catalyzed reaction to occur.

One of the most fascinating questions in biology is how organs, cells, organelles, filaments and other macromolecular complexes, which are often constructed from a series of repeating units, reach a defined size and then stop growing. The control of the length of the cylindrical coat of TMV is a good example (Example 5.3). Here, it is clear that the interaction between the coat protein and the RNA which it encloses is critical. In other cases, such as the length of the bacteriophage T4 tail

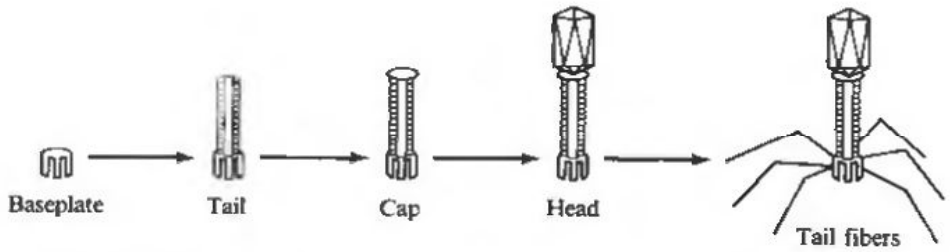


Fig. 5-3 The stepwise assembly sequence of the bacteriophage T4 virus.

Fig. 5-4 An example of the vernier principle.

and the highly regulated length of the repeating contractile units within skeletal muscle fibers, the answers are not so clear.

There are several possible explanations, but one simple mechanism based on the vernier principle is as follows. Two rod-like proteins of different lengths aggregate in parallel to form a linear complex (Fig. 5-4). This will grow until the ends of each rod exactly coincide. Now that there is no overlapping segment the complex will grow no further.

5.3 PROTEIN SELF-ASSOCIATION

A large number of biologically active proteins exist in solution as large complexes. The simplest cases of self-assembly involve the *self-association* of a single type of subunit. Many proteins possess a quaternary structure in which identical subunits are assembled into geometrically regular structures: e.g., the storage form of insulin contains six monomers arranged in the shape of a hexagon, while the iron-storage protein ferritin contains 24 subunits arranged as an icosohedron. Some aspects of symmetry and self-association are described below.

Symmetrical Dimers

Question: In what ways can proteins self-associate to form geometrically regular oligomers?

If the binding site is complementary to itself, then a *symmetrical dimer* will be formed [Fig. 5-5(*b*)]. There will be a *diad* (or twofold rotational) axis of symmetry between the two subunits, such that a rotation of one subunit by 180° about this axis will superimpose it onto the other subunit. (You should prove to yourself that a diad axis of symmetry cannot be formed if the binding site is not complementary to itself.) The dimer so formed may itself act as a subunit of larger aggregates; e.g., two dimers may associate through a different binding interface to generate a tetramer, with two axes of symmetry. In such cases, where two different axes of symmetry exist, the symmetry is described as *dihedral*.

The term *protomer* is often used to describe the basic unit taking part in a self-association reaction. In the above question, the *monomer* is the protomer of the dimer. Similarly, since two identical dimers associate to form the tetramer, hexamer and so on, the dimer could be described as the protomer of the second association [Fig. 5-5(*c*)].

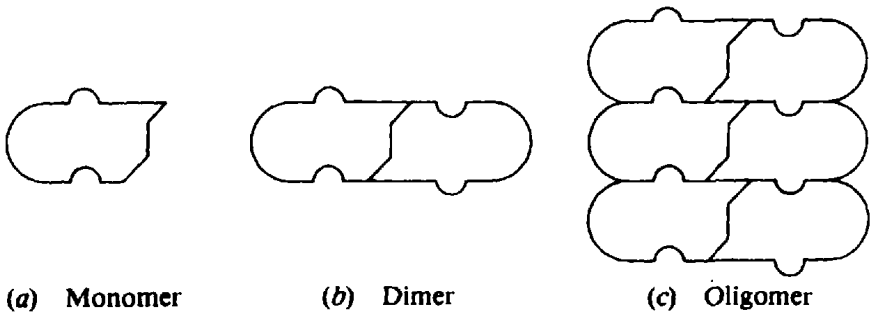


Fig. 5-5 Examples of simple self-association of a protomer.

Rotational Symmetry

If the binding site is complementary to a site elsewhere on the protein, then a *chain* will be formed. For certain angles between the protomers, the chain will close upon itself and form a *ring* (Fig. 5-6). This ring may vary in size from a dimer to larger oligomers. Such a regular ring possesses *rotational symmetry* and this type of symmetry is commonly found in proteins having three, five or other uneven numbers of protomers, although it may also be found among oligomers with an even number of protomers (such as the storage form of insulin, which is a hexamer).

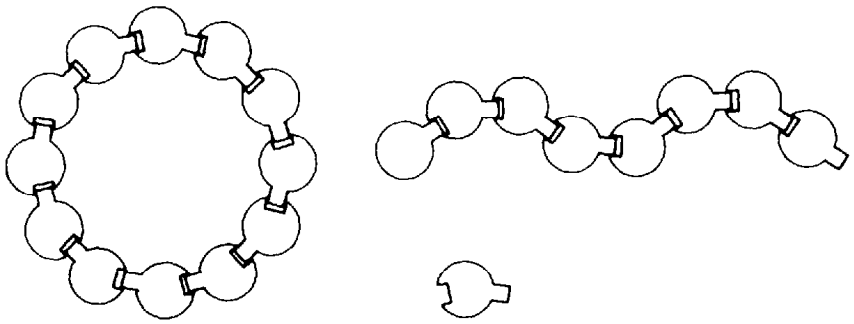


Fig. 5-6 Rings or helices can be formed when protein subunits interact with each other at appropriate fixed angles.

EXAMPLE 5.5

Hemagglutinin, found in the membrane of the influenza virus, is a *trimer* (Fig. 5-7). Each protomer is a single polypeptide chain folded into two segments: an α-helical "stalk", and a globular head containing an eight-stranded β-sheet structure which contains the binding site for receptors on the host cells. The "stalks" from each protomer interact through their nonpolar amino acid residues to form a left-handed superhelix, or coiled coil. These hydrophobic interactions provide the major forces for stabilizing the trimer. The resulting threefold cyclic axis of symmetry is shown in Fig. 5-7.

Indefinite Self-Association

In many cases, a chain can form which is *open-ended*, and additional protomers can continue to be added to the free binding sites at the ends of the chain without limit. This is termed an *indefinite* self-association. If the angle between the protomers is fixed, the open-ended chain will be in the form a helix (Fig. 5-6).

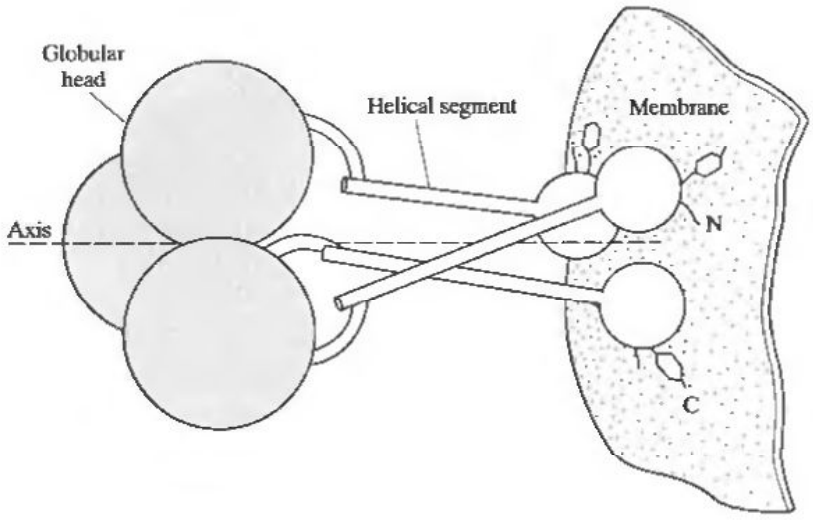


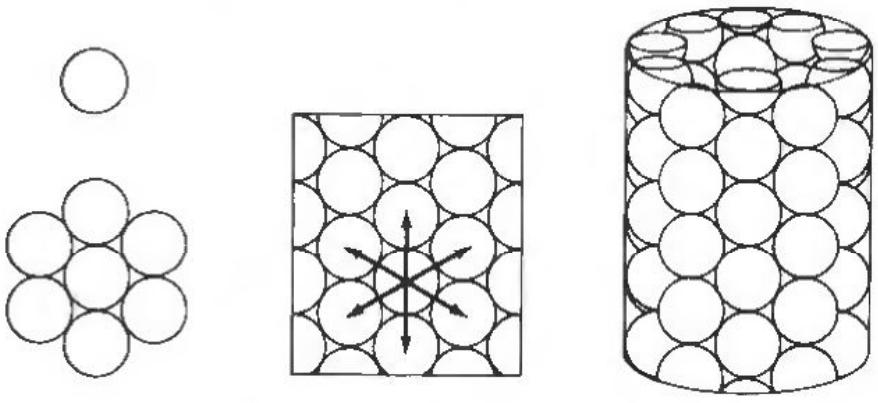
Fig. 5-7 A diagrammatic representation of the hemagglutinin trimer from the influenza virus.

If two helical strands are wound around each other to form a double helix, or if monomers in succeeding turns of the helix are in contact, there is a considerable increase in stability since each monomer interacts with two monomers in the opposite strand as well as with its two neighbors in its own strand [Fig. 5-8(*a*)].

Sheets and Closed Surfaces

Multiple interactions in the same plane can lead to the formation of sheets where, for example, each monomer can interact with six neighbors in a hexagonal close-packing arrangement (Fig. 5-8). Sheets can, with a slight readjustment, be converted into cylindrical tubes (Fig. 5-8) or even into spheres. These closed structures can provide even greater stability since they maximize the number of interactions that can be made. The protein coats of certain viruses are excellent examples of this. *Microtubules*, which consist of the protein *tubulin*, can be converted readily between sheet and tubular forms, at least in the purified form.

(*a*) Double helix, e.g., actin



(*b*) Sheet (*c*) Tube

Fig. 5-8 The formation of double helices, sheets and tubes, showing multiple interactions (arrows) between subunits.

EXAMPLE 5.6

In *tomato bushy stunt virus*, 180 identical protein subunits (M_r = 41,000) form a shell which surrounds a molecule of RNA containing 4,800 nucleotides. For geometrical reasons no more than 60 identical subunits can be positioned in a spherical shell in a precisely symmetrical way. This limits the volume, so, in order to accommodate a greater amount of RNA, a larger shell is needed. This can be achieved by *relaxing* the symmetry: the subunits are divided into three sets of 60, each set packing with strict symmetry. However, the relationship between each set is different so that the overall packing is only quasi-equivalent.

Equilibria

Association reactions can be characterized by *equilibrium constants*. Experimental determination of equilibrium constants for each step in an association reaction provides vital information about the properties of the associating system. In particular, the *mode* of association (e.g., monomer–dimer, monomer–tetramer, indefinite), and the *strength* of the association (that is, the degree to which various oligomers can exist at various total concentrations) can be obtained. The evaluation of equilibrium constants over a range of solution conditions (such as salt concentration and temperature) can be used to obtain information on the *enthalpy* and *entropy* of the various steps in the association and the *types of bonds* involved in the assembly process. Note that this information can be obtained in the complete absence of structural information, although, of course, any available structural information can be used to aid in the interpretation of the thermodynamic data.

EXAMPLE 5.7

The dimerization reaction of a *solute*:

$$2A \rightleftharpoons A_2$$

can be characterized by a *dimerization association constant*,

$$K_2 = [A_2]/[A]^2 \qquad (5.1)$$

where the square brackets denote molar concentration. (This definition of the equilibrium constant is valid at low concentrations of solute; that is, where the system is behaving *ideally*.) The relationship shows that the proportion of dimer increases with the total concentration of the molecule. Conversely, dilution favors *dissociation* (see Prob. 5.4).

Question: Are the properties of an associating system observed *in vitro* always the same as the properties of the same system *in vivo*?

Not necessarily! Experiments *in vitro* are usually performed under nearly ideal conditions. On the other hand, *in vivo* conditions (such as the cytoplasm, or the interior of an organelle) are extremely *crowded*. The total concentration of protein is usually in the range 200–500 g L^{-1}, and high concentrations of small molecules are also present. Up to 50 percent or more of the aqueous environment can be occupied by all of these solutes. These crowded environments are highly *nonideal*, and the association properties of the system can be grossly altered as a result.

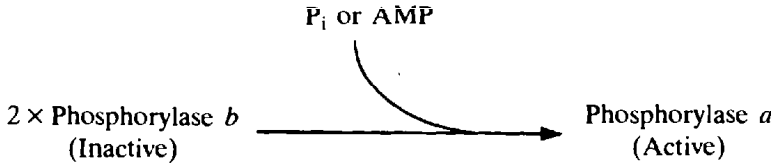
If, for example, the total volume occupied by a dimer is less than the total volume occupied by two monomers, then the dimer will become more and more favored as the aqueous environment becomes more and more crowded. The changes can be dramatic, with the association being barely detectable under ideal conditions, but *strongly* favored under the crowded conditions found in a cell.

Catalysis by enzymes can also be affected by crowding in a number of ways. For example, an enzyme might undergo a monomer-tetramer reaction where only the tetramer is catalytically active. If tetramer formation is favored under crowded conditions then the rate of reaction will increase. In this way, rate constants can be altered by several orders of magnitude.

The binding of other molecules (small or large) may also change the degree of association. For example, if the associated form binds a small molecule (or ligand) better than the protomer, then the presence of that ligand will promote association. Conversely, if the ligand binds preferentially to the protomer then dissociation is promoted. This provides one crucial means of regulating the polymerization state, assembly and disassembly of supramolecular structures.

EXAMPLE 5.8

Rabbit muscle phosphorylase can exist in two forms; an essentially inactive dimer, phosphorylase *b*, and an active tetramer, phosphorylase *a*. When AMP is noncovalently bound to phosphorylase *b*, or when phosphorylase *b* is phosphorylated at a serine residue by *phosphorylase kinase*, the enzyme is converted to the active, predominantly tetrameric, form (Chap. 11). The reaction can be reversed by the removal of AMP or the *dephosphorylation* of serines by *phosphorylase phosphatase*.



This example not only highlights the effect of noncovalent ligand binding on association (and catalysis) but also the role played by covalent phosphorylation and dephosphorylation of selected residues (serine, threonine and tyrosine) in proteins. Phosphorylation/dephosphorylation is an extremely important regulator of the assembly, disassembly and other dynamic properties of many supramolecular structures.

Heterogeneous Association

Most supramolecular structures are *heterogeneous*. That is, they contain more than one type of subunit. Frequently the different subunits have different functional roles, e.g., *catalytic*, *regulatory*, or purely structural.

EXAMPLE 5.9

The enzyme *aspartate transcarbamoylase* catalyzes an early regulated step in the synthesis of pyrimidine nucleotides (Chap. 15). The enzyme from *E. coli* can be dissociated into two kinds of subunit, a *catalytic* subunit, C_3 (M_r = 100,000), and a *regulatory* subunit, R_2 (M_r = 34,000). The catalytic subunit is a trimer and is catalytically active, but is not regulated. The regulatory subunit is a dimer. It has no catalytic activity but will bind the regulators ATP and CTP. When mixed, the subunits self-assemble, the resulting complex being fully active and regulated by the two nucleotides:

$$2C_3 + 3R_2 \rightleftharpoons R_6C_6$$

x-Ray diffraction studies show the two catalytic trimers lying back-to-back with the three regulatory dimers fitting into grooves on the outside (Fig. 5-9). The catalytic (active) sites are in the center, near the threefold rotational axis of symmetry, while the regulatory binding sites are on the outside, far from the active sites. The binding of a regulatory molecule (ATP or CTP) causes a conformational change which is transmitted across the complex to the active site.

In some instances, the specificity of an enzyme may be altered by its association with a *modifier* subunit. An example is *lactose synthase*, the enzyme that catalyzes the linking of glucose and galactose to form lactose. This enzyme consists of a catalytic subunit and a modifier subunit. By itself, the catalytic subunit catalyzes the addition of a galactose residue to a carbohydrate chain of a glycoprotein; in the presence of the modifier (α-lactalbumin) the specificity is changed and lactose is synthesized.

It can be advantageous for a series of enzymes catalyzing a sequence of reactions in a metabolic pathway to assemble into a *multienzyme complex*. This assembly increases the efficiency of the pathway in that the product of one enzymatic reaction is in place to be the substrate of the next

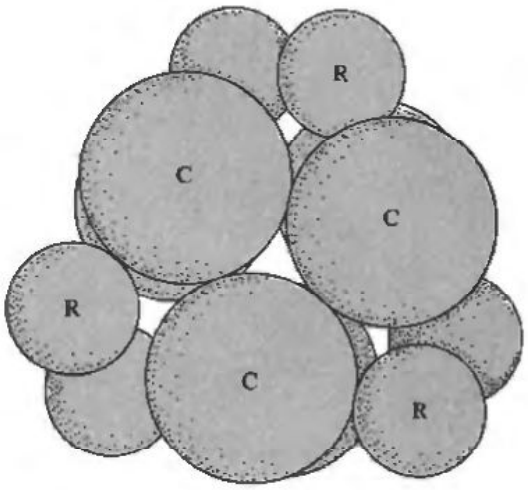


Fig. 5-9 The arrangement of the catalytic (C) and regulatory (R) subunits in aspartate transcarbamoylase from *E. coli*.

(Chap. 9). The limitation imposed by the rate of diffusion of the reactants in solution is therefore largely overcome.

EXAMPLE 5.10

The *pyruvate dehydrogenase complex* catalyzes the conversion of pyruvate to acetyl-CoA. This conversion links the breakdown of carbohydrates to the processes of respiration and oxidative phosphorylation (Chap. 12). The *overall* reaction is:

$$\text{Pyruvate} + \text{CoA} + \text{NAD}^+ \rightleftharpoons \text{Acetyl-CoA} + \text{CO}_2 + \text{NADH} + \text{H}^+$$

However, this reaction is in fact the sum of five reactions catalyzed by three enzymes and requiring five cofactors: coenzyme A (CoA), NAD^+, FAD, thiamine pyrophosphate (TPP), and lipoic acid (an 8-carbon acid with sulfhydryl groups at positions 6 and 8). The detailed steps of these reactions are shown in Fig. 12-7.

The lipoic acid is covalently attached to the *transacetylase* (E_2). It acts as a 'swinging arm', interacting with pyruvate decarboxylase (E_1) to accept the hydroxyethyl derivative of pyruvate, with CoA to produce acetyl-CoA, and with *dihydrolipoyl dehydrogenase* (E_3) so that it can be reoxidized.

In *E. coli*, the complex has a mass of about 4×10^6 Da and consists of 60 polypeptide chains. In the center of the complex there is a core of eight trimers of E_2 arranged in cubic symmetry (Fig. 5-10). Dimers of E_3 are bound to the six faces of the cube. Finally, pairs of E_1 bind to each edge of the cube encircling the dehydrogenase dimers. The central position of the transacetylase (E_2) allows the flexible lipoyl 'arms' to transfer reactants from E_1 to E_3 or to CoA.

The mammalian enzyme complex is even larger, containing almost 200 polypeptide chains to give a molecular weight of over 7×10^6. As well as the enzymatic activities described above, the mammalian complex contains two further enzymes which act as regulators by catalyzing phosphorylation/dephosphorylation of pyruvate decarboxylase in response to the metabolic demand: the complex is activated by dephosphorylation when there is a need for the product acetyl-CoA and inactivated by phosphorylation when acetyl-CoA is not required.

Enzyme complexes performing similar or identical tasks can vary widely between species. An excellent example is the enzyme complex, fatty acid synthase, which catalyzes the synthesis of fatty acids from acetyl-CoA and involves seven catalytic steps (Chap. 13). In *E. coli* and most bacteria the complex consists of seven different enzymes. In more advanced bacteria and in eukaryotic cells there are fewer types of subunit. For example, the yeast enzyme is a multienzyme complex ($M_r = 2.3 \times 10^6$) with just two types of subunit (A and B) and a stoichiometry of A_6B_6. The subunits are *multicatalytic*. Subunit A ($M_r = 185{,}000$) has three catalytic activities and subunit B ($M_r = 175{,}000$) has the remaining four. The mammalian liver complex is a dimer, with each subunit

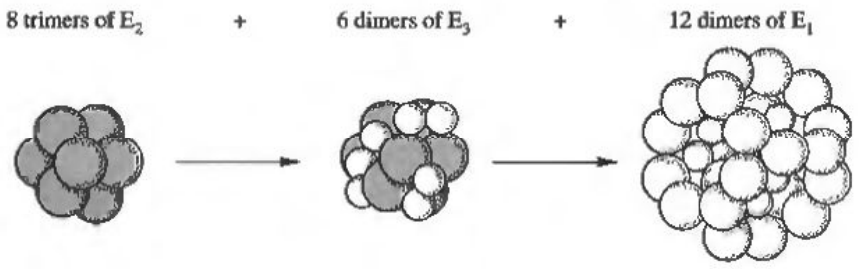


Fig. 5-10 Model of the *E. coli* pyruvate dehydrogenase complex.

consisting of a single polypeptide chain ($M_r = 240{,}000$) containing all the enzymatic activities necessary for fatty acid synthesis.

5.4 HEMOGLOBIN

The ability to control biological activity can be enhanced by the formation of complexes (Examples 5.8–5.10). The best documented example of this is hemoglobin. Hemoglobin is a *tetramer* consisting of two α and two β chains (Chap. 4). The chains are similar in structure to each other and to myoglobin. Each of the four chains folds into eight α-helical segments (the *globin fold*) designated A to H from the N-terminus.

Quaternary Structure

There are two types of contact between the hemoglobin chains. The contacts within both the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers involve the B, G and H helices and the GH loop (Fig. 5-11). These are known as the *packing contacts*. The contacts between these two dimers (α_1 with β_2 and α_2 with β_1) involve

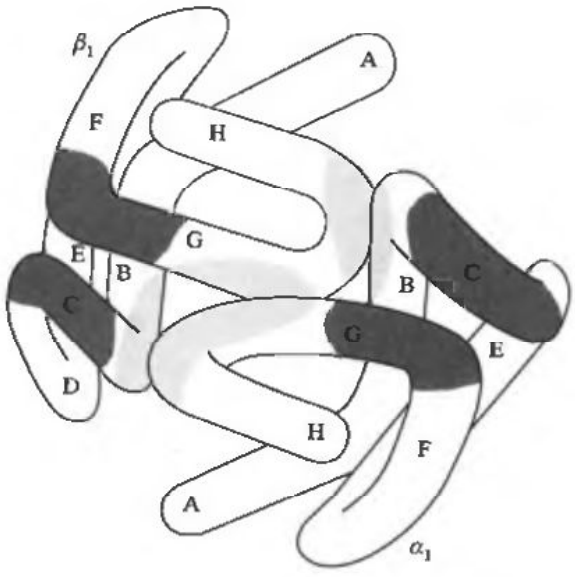


Fig. 5-11 A view of the $\alpha_1\beta_1$ dimer of hemoglobin. The packing contacts (light gray) hold the dimer together. The sliding contacts (dark gray) form interactions with the $\alpha_2\beta_2$ dimer.

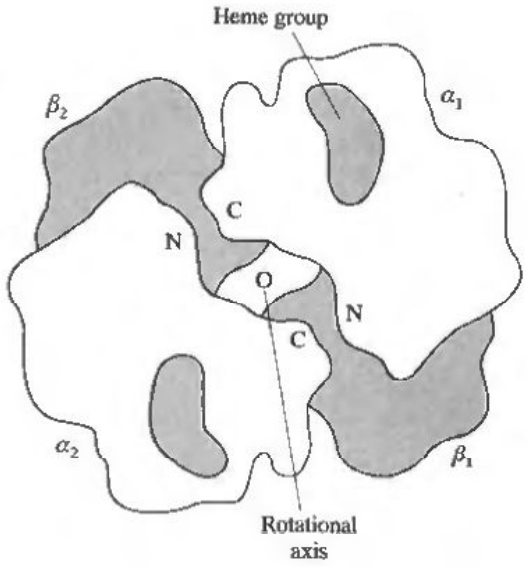


Fig. 5-12 Diagrammatic representation of hemoglobin showing its cyclic symmetry.

the C and G helices and the FG loop (Fig. 5-11). These are called the *sliding contacts* because movement between dimers can occur here. About 20% of the surface area of the chains is buried in forming the tetramer. These contacts are mainly hydrophobic with about one third of the contacts involving polar side chains in hydrogen bonds and electrostatic interactions (or salt links).

The tetramer of hemoglobin may be considered to be a *symmetrical* molecule made up of two *asymmetrical* but identical protomers, the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers. They are related by a twofold rotational axis of symmetry: if one dimer is rotated by 180° it will superimpose on the other (Fig. 5-12).

If the α and β chains are considered to be identical, then hemoglobin has dihedral symmetry with two rotational axes, and with the four subunits arranged at the apices of the tetrahedron.

Question: Why do hemoglobin chains associate while myoglobin chains do not?

Although the sequences of the myoglobin and hemoglobin chains are *homologous* and they all adopt the globin fold, there are important differences in the regions where the chains in hemoglobin contact one another. In particular, some of the surface residues which are polar in myoglobin are hydrophobic in hemoglobin. For example, residue B15 (the 15th residue in the B helix) is lysine in myoglobin and leucine or valine in hemoglobin, and residue FG2 (the second residue in the loop between the F and G helices) is histidine in myoglobin and leucine in both chains of hemoglobin. In hemoglobin, both these residues are part of the contact between the α and β chains.

Oxygen Binding and Allosteric Behavior

When hemoglobin binds oxygen to form *oxyhemoglobin* there is a change in the quaternary structure. The $\alpha_1\beta_1$ dimer rotates by 15° upon the $\alpha_2\beta_2$ dimer, sliding upon the $\alpha_1\beta_2$ and $\alpha_2\beta_1$ contacts, and the two β chains come closer together by 0.7 nm.

The change in quaternary structure is associated with changes in the tertiary structure triggered by the movement of the proximal histidine (the histidine nearest to the heme iron) when the heme

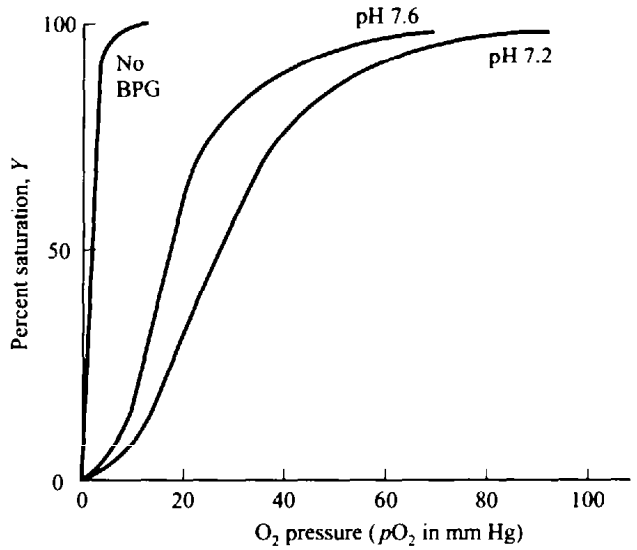
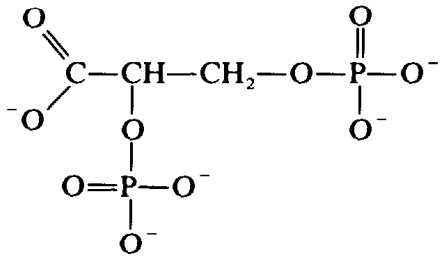


Fig. 5-13 Effect of BPG and pH on the oxygen affinity of hemoglobin.

iron binds oxygen. These changes cause the breakage of the constraining salt links between the terminal groups of the four chains. Deoxyhemoglobin is thus a structure with a low affinity for oxygen (the so-called *tense structure*) while oxyhemoglobin has a structure with a high affinity for oxygen (the so-called *relaxed structure*). Thus, as oxygen successively binds to the four heme groups of the hemoglobin molecule, the oxygen affinity of the remaining heme groups increases. This cooperative effect (Chap. 9) produces a sigmoidal oxygen dissociation curve (Fig. 5-13).

BPG

The compound *2,3-bisphosphoglycerate* (BPG, also known as 2,3-diphosphoglycerate or DPG) is produced within the red blood cell of many animal species, and acts to modify the oxygen binding affinity of hemoglobin:



2,3-Bisphosphoglycerate (BPG)

This polyanionic compound carries five negative charges under physiological conditions and binds in the central cavity (the β-β cleft) of *deoxy*hemoglobin, making salt links with cationic groups on the β chains (Fig. 5-14). In oxyhemoglobin, the cavity is too small to accept BPG. The binding of oxygen and BPG are thus mutually exclusive, with the effect that BPG reduces the oxygen affinity of hemoglobin. (Compare, in Fig. 5-13, the dissociation curve at pH 7.6, where BPG is present, with that labeled "No BPG".)

Question: At high altitudes the concentration of BPG in the red blood cell can increase by 50% (after 2 days at 4,000 m). How does this assist adaptation to a lower partial pressure of oxygen?

As the concentration of BPG increases, the affinity of oxygen for hemoglobin decreases and more oxygen is released from the blood to the tissues. This compensates for the decreased arterial oxygen concentration.

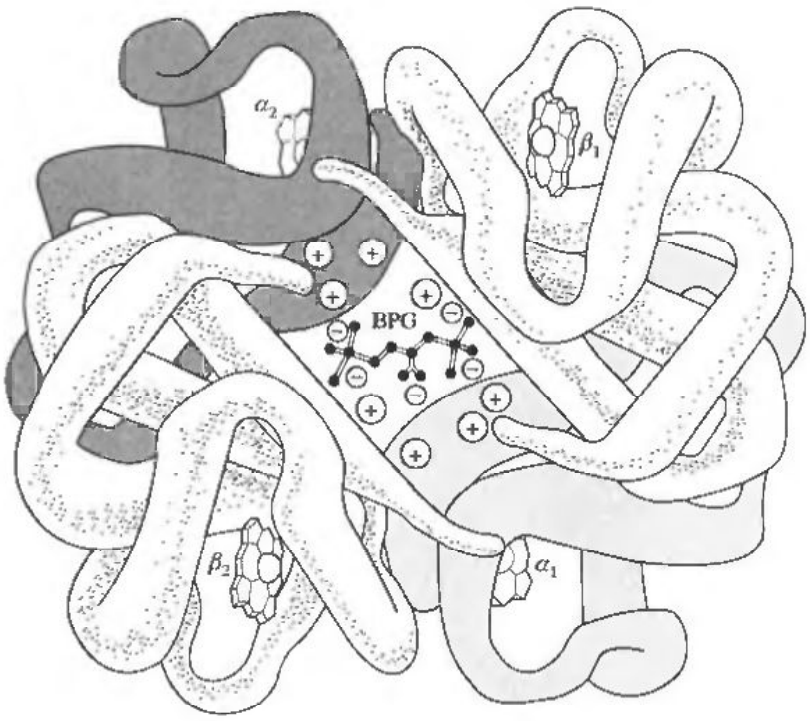


Fig. 5-14 Binding of 2,3-BPG in the β-β cleft of deoxyhemoglobin.

pH and CO_2

Rapidly metabolizing tissues require increased levels of oxygen, and therefore there is a requirement for oxyhemoglobin to release more oxygen to the cells of these tissues. In rapidly metabolizing tissues there is a fast build-up of CO_2 from the oxidation of fuels such as glucose. This causes an increase in proton concentration (decrease in pH) through the following reaction:

$$H_2O + CO_2 \rightleftharpoons HCO_3^- + H^+$$

Deoxyhemoglobin has a higher affinity for protons than does oxyhemoglobin, so that the binding of protons competes with the binding of oxygen (though at different sites):

$$Hb(O_2)_4 + 2H^+ \rightleftharpoons Hb(H^+)_2 + 4O_2$$

This effect, known as the *Bohr effect*, arises from the slightly higher pK_a of ionizing groups in deoxyhemoglobin. One such group is histidine β146 in which the pK_a is raised as a consequence of its proximity to a neighboring aspartate in deoxyhemoglobin and which therefore has a higher affinity for protons. In oxyhemoglobin, as a result of a changed geometry, histidine β146 is free and has a more normal pK_a. A decrease in pH from 7.6 to 7.2 can almost double the amount of oxygen released in the tissues (see Fig. 5-13). In the lungs, where the oxygen tension is very high, the two reactions shown above are partly reversed, leading to the release of CO_2.

Although much of the CO_2 in blood is transported as HCO_3^-, some combines with hemoglobin and acts as an *allosteric effector* (Chap. 9). CO_2 reacts with the unionized form of α-amino groups

in deoxyhemoglobin to form *carbamates* which can form salt links, thus stabilizing the tense structure:

$$\text{Hb-NH}_2 + \text{CO}_2 \rightleftharpoons \text{Hb-NHCOO}^- + \text{H}^+$$

In conclusion, the interactions between the subunits of hemoglobin allow the release of oxygen to be fine-tuned to physiological needs. The allosteric effectors BPG, H^+ and CO_2 all lower the affinity of hemoglobin for oxygen by increasing the strength of the subunit interactions.

5.5 THE EXTRACELLULAR MATRIX

In multicellular organisms, the space between cells is largely occupied by an intricate network of interacting macromolecules that constitute the *extracellular matrix*. This matrix surrounds, supports and regulates the behavior of the cells, such as *fibroblasts*, that secrete the extracellular matrix. Together, the matrix with these cells constitutes the *connective tissue*.

The extracellular matrix is made up of fibrous proteins, such as *collagen* and *elastin*, embedded in a gel-like ground substance of *proteoglycan* and water. Other proteins, such as *fibronectin*, play a specialized role in binding the matrix components together. In calcified tissue, a *filler* of *hydroxyapatite* crystals, $Ca_{10}(PO_4)_6(OH)_2$, replaces the ground substance to enable high load bearing. Both the amount of connective tissue and the composition of the macromolecular components vary greatly from one organ to another. This gives rise to an enormous diversity of forms, functions and physical properties (compare marble-like bones and teeth to rope-like tendons; see Table 5.1).

Table 5.1. Approximate Composition of Connective Tissues (% of Dry Weight*)

	Hydroxyapatite	Collagen	Elastin	Proteoglycan
Bone	80	19	0	1
Cartilage	0	50	0	50
Tendon	0	90	9	1
Skin	0	50	5	45
Ligament	0	23	75	2

*The water content of bone is low (~10%), but high in the others (~70%).

Collagen

Structure

The *collagens* are a widely distributed family of proteins; in mammals they make up over 25 percent of the total body protein. The typical collagen molecule consists of a stiff, inextensible, triple-stranded helix (Chap. 4). Twenty five genetically different individual polypeptide chains have now been described, and these combine in triplets to form 15 collagen types (Table 5.2). Of these, type I, the most abundant, has been the most fully investigated. Each of its three α chains, with a mass of 95,000 Da, contains about 1,000 amino acids, of which glycine constitutes one-third and proline one-fifth. Early investigators denatured collagen to form gelatin and found molecular species of approximately 100,000, 200,000, and 300,000 Da, which they called α, β, and γ chains. In due course, it was discovered that the β chains were dimers and the γ chains trimers of the α chains, but the nomenclature has remained.

Table 5.2. Types and Properties of Collagen

Type	Chain Designations, Approximate Triplet Composition and Length	Structural Form	Distribution
I	2 α1(I) and 1 α2(I); 300 nm	Broad fibrils with 67-nm bands	Most abundant; found in skin, bone, tendon, and cornea
II	3 α1(II); 300 nm	Small fibrils with 67-nm bands	Abundant, in cartilage, vitreous humor, and intervertebral disks
III	3 α1(III); 300 nm	Small reticulin fibrils with 67-nm bands	Abundant; found in skin, blood vessels, and internal organs
IV	2 α1(IV), 1 α2(IV); 390 nm	Nonfibrillar cross-linked network	Found in all basement membranes (a delicate layer under the epithelium of many organs)
V	α1(V), α2(V), α3(V); 390 nm	Small fibers, often associated with type I	Widespread in interstitial tissues
VI	α1(VI), α2(VI), α3(VI); 150 nm	Fibrils with 100-nm bands, associated with type I	In most interstitial tissues
VII	3 α1(VII); 450 nm	End-to-end dimers that assemble laterally	Anchoring fibrils between basement membrane and stroma
IX	α1(IX), α2(IX), α3(IX) 200 nm	Laterally associated with type II, bound glycosaminoglycans	Minor cartilage protein
X	3 α1(X); 150 nm	Form sheets	In hypertrophic or mineralizing cartilage

Question: By what processes do three single α chains line up to form a regular triple helix?

The individual chains are synthesized as *procollagen* α chains, which, with a mass of 150,000 Da, have additional *extension peptides* at both the amino and carboxyl termini. The amino- and carboxyl-terminal regions from the three α chains each fold to form globular structures, which then interact to guide the formation of the triple helix. Interchain disulfide bonds stabilize the carboxyl-terminal domain and the triple helix coils up from this domain (Fig. 5-15).

Question: Collagen is unusual in that it contains residues of both hydroxyproline and hydroxylysine. As these amino acids cannot be directly incorporated into the polypeptide chain, how are they formed in collagen?

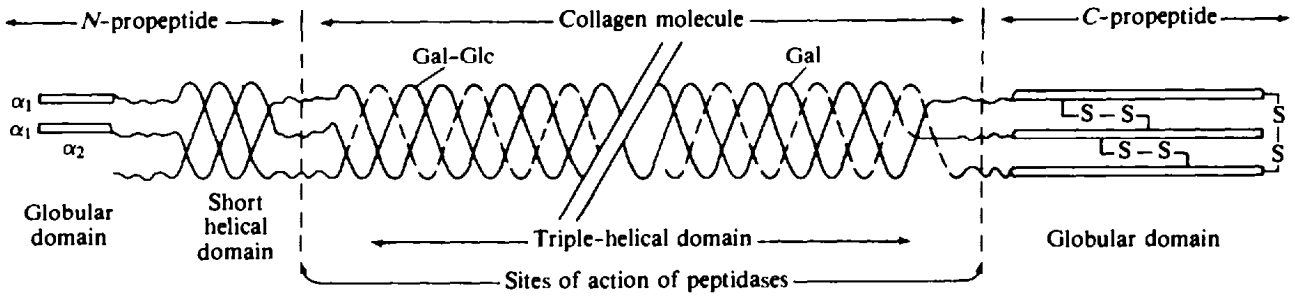


Fig. 5-15 Model of type I procollagen.

The enzyme *prolylhydroxylase* catalyzes the hydroxylation of proline residues, which lie immediately before the repetitive glycine residues in the pro-α-chains:

$$\text{—Prolylglycyl—} + \text{2-Oxoglutarate} + O_2 \xrightarrow[\text{Ascorbate}]{Fe^{2+}} \text{—Hydroxyprolylglycyl—} + \text{Succinate} + CO_2 + H^+$$

The enzyme *lysylhydroxylase* acts in a similar way on lysine residues. Note that the triple-helical portion of collagen consists of $(x\text{-}y\text{-Gly})_n$, where x is often proline and y is commonly hydroxyproline.

Question: In scurvy, caused by a lack of ascorbate (vitamin C), the skin and the blood vessels become extremely fragile. Why is that?

Without ascorbate, hydroxylation of proline and lysine residues does not take place. This has three effects: (1) It prevents the formation of interchain hydrogen bonds involving the hydroxyl groups of hydroxyproline, and thus the triple helix is less stable. (2) It prevents *glycosylation*, which is the addition of galactose units to hydroxylysine residues, followed in most cases by the addition of a glucose to the galactose. (3) It limits the extent of cross-linking of mature collagen molecules (see Fig. 5-18). Most of the nonhydroxylated pro-α-chains are degraded within the cell.

Question: Once the procollagen molecule is hydroxylated, glycosylated, and coiled into a triple helix, how is it processed to form higher structures?

The procollagen molecule is secreted from the cell, and the extension propeptides are excised by two specific *procollagen peptidases* to form the *tropocollagen* molecule. The removal of the peptides (M_r = 20,000 and 35,000) allows the tropocollagen molecules to self-assemble to form *fibrils*. This assembly is regulated to some extent by the cells and by other extracellular components to produce the wide variety of structures found in collagen fibers.

For electron microscopy, collagen fibers are fixed and stained with heavy-metal reagents; they show cross-striations of dark and light bands every 67 nm (Fig. 5-16). This effect arises from accumulation of the heavy metal in the gap regions to produce the dark-stained bands. The individual tropocollagen molecules assemble in such a way that adjacent molecules are staggered, i.e., displaced longitudinally by approximately one-quarter of their length (67 nm) and leaving a gap between the ends of each molecule.

This packing of the tropocollagen molecules with a displacement of 67 nm is caused by repeating clusters of charged and uncharged residues along the polypeptide chains with a periodicity of 67 nm. Hence the maximum number of intermolecular interactions (electrostatic and hydrophobic) is formed when the tropocollagen molecules are displaced by multiples of 67 nm.

However, it is more difficult to explain how the tropocollagen molecules pack three-dimensionally in the quarter-staggered array to produce a cylindrical fibril. One favored model is a pentamicrofibril consisting of groups of five tropocollagen molecules packing together in the 67-nm staggered pattern (Fig. 5-17).

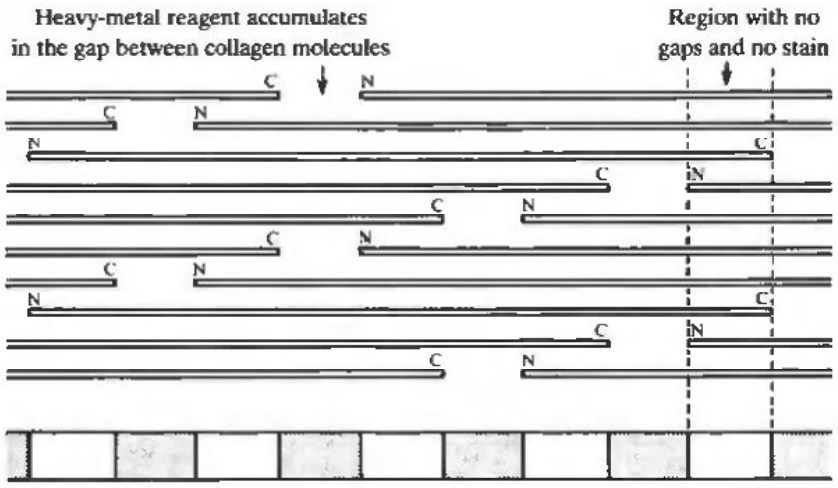


Fig. 5-16 Schematic diagram showing how the staggered arrangement of tropocollagen molecules gives rise to cross-striations in a fibril negatively stained with phosphotungstic acid.

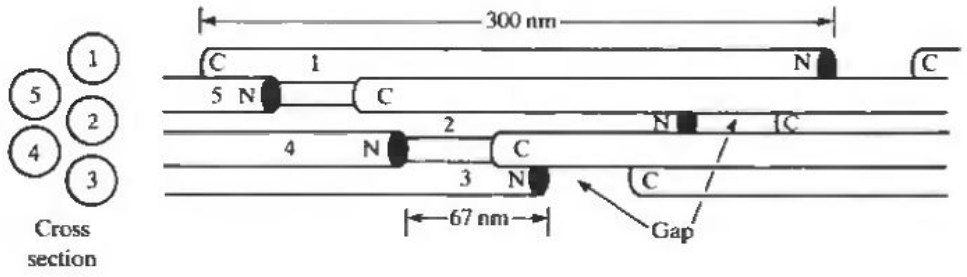


Fig. 5-17 A possible model for the three-dimensional packing arrangement of collagen molecules in a pentamicrofibril.

The size and arrangement of the collagen fibrils vary considerably from tissue to tissue, the diameter ranging from 10 to 300 nm and the packing from being apparently random, as in mammalian skin, to being in strict parallel bundles, as in tendon.

Cross-Linking of Collagen

The collagen fibrils are further strengthened to withstand high tensile forces through the introduction of cross-links. Covalent cross-links are formed both within a tropocollagen molecule and between different molecules. The first step is the enzymatic *oxidative deamination* of the ε-amino group of lysine residues to form the aldehyde group of *allysine* residues. These highly reactive aldehydes then spontaneously react with each other or with lysine residues (Fig. 5-18) to form covalent bonds, which can often react further with other residues such as histidine.

Collagen types I, II, III, V and XI form fibrillar structures. The others often contain nonhelical regions but all aggregate to form a variety of supramolecular structures, with a variety of forms (Table 5.2).

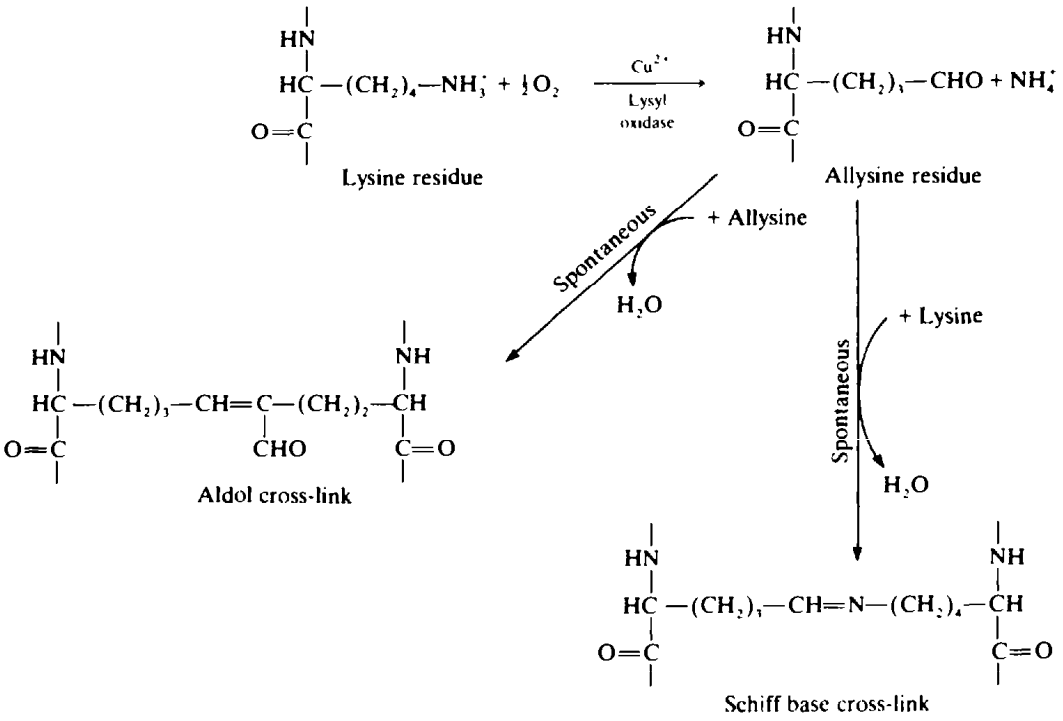


Fig. 5-18 Reactions that produce cross-links in collagen.

EXAMPLE 5.11

The secreted form of type IV collagen contains a globular region at the C terminus and a bend in the triple helix near the N terminus. It polymerizes without any further processing and cannot associate laterally, but it still polymerizes to form a two-dimensional sheetlike network. The asymmetric monomers aggregate only through mutual association of their identical ends; the globular regions of two monomers bind together, while four monomers are held together by their triple-helical N termini (Fig. 5-19).

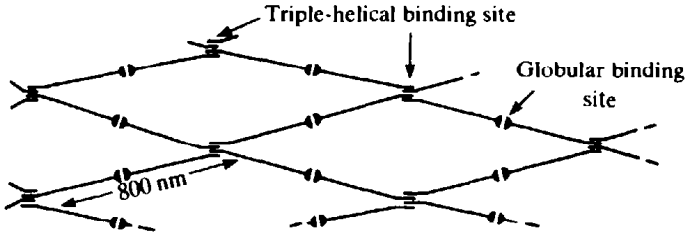
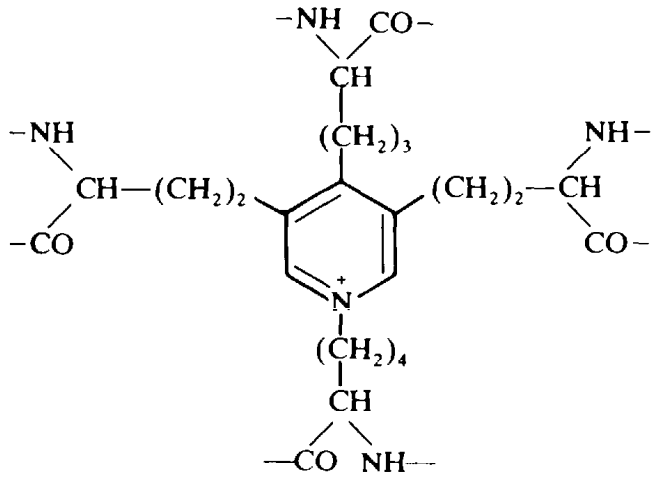


Fig. 5-19 A stylized model of the association of type IV collagen in basement membrane.

Elastin

The other major protein in the extracellular matrix is *elastin*, which is the main component of elastic fibers found in ligaments, large arteries, and lungs. After synthesis and partial hydroxylation of proline residues, a 72 kDa molecule of *tropoelastin* is secreted into the matrix. This protein is rich in nonpolar amino acids and contains repeating sequences, such as (Val-Pro-Gly-Val-Gly). These sections form an amorphous, random-coiled structure with frequent *reverse turns*. Other recurrent sequences are rich in alanine with paired lysine residues, e.g., -Ala-Ala-Ala-Ala-*Lys*-Ala-Ala-*Lys*-

Ala. The action of lysyl oxidase to produce allysine allows three of these modified residues to condense with one lysine residue to form the heterocyclic complex amino acid *desmosine*, which cross-links two or even three chains. A highly cross-linked network results.



Desmosine residue

EXAMPLE 5.12

Elastin is not a true *rubber* as it is not self-lubricating. It has elastic properties only in the presence of water. At rest, elastin is tightly folded, stabilized by hydrophobic interactions between nonpolar residues; this has been termed an *oiled coil*. On stretching, these hydrophobic interactions are broken, and the nonpolar residues are exposed to water. This conformation is thermodynamically unstable, and once the stretching force is removed, the elastin recoils to its resting state.

The amorphous elastin is organized into *elastic fibers* with a sheath of *microfibrils* surrounding each 10 nm fiber. The major component of the sheath is the glycoprotein *fibrillin*, mutations in whose gene causes *Marfan syndrome*.

Proteoglycans

The ground substance of the extracellular matrix is a highly hydrated gel containing large polyanionic *proteoglyan* molecules, which are about 95 percent polysaccharide and 5 percent protein. The polysaccharide chains are made up of repeating disaccharide units and are called *glycosaminoglycans* because one half of the disaccharide is always an amino sugar derivative, either *N*-acetylglucosamine or *N*-acetylgalactosamine; the other half is usually a uronic acid, such as glucuronic acid, giving a negative charge (Chap. 2).

EXAMPLE 5.13

Hyaluronan is a single, very long glycosaminoglycan chain having from 500 to several thousand repeating disaccharide units, e.g., $[\beta\text{-}(1\rightarrow4)\text{-GlcA-}\beta\text{-}(1\rightarrow3)\text{-GlcNAc-}\beta\text{-}(1\rightarrow4)]_n$ (see Fig. 5-20). These molecules have molecular weights between 0.2×10^6 and 10×10^6 and may exist without being bound covalently to a protein. Because of the carboxyl groups, this substance carries a large negative charge at neutral pH values. It is often called hyaluronic acid, or, more appropriately, hyaluronate.

The *core protein* of the proteoglycan is unusually large ($M_r \simeq 300{,}000$), with a globular head ($M_r \simeq 75{,}000$) and a long tail rich in serine and threonine residues. The glycosaminoglycans are covalently attached to these hydroxyl amino acids via an *oligosaccharide linkage* (see Fig. 5-21). During synthesis, the monosaccharides are added one at a time by a specific *glycosyltransferase*;

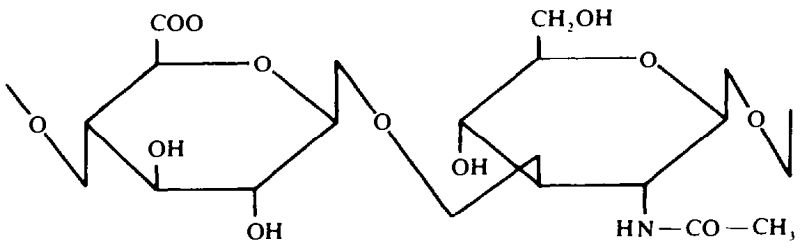


Fig. 5-20 Structural formula of the repeating disaccharide of hyaluronan.

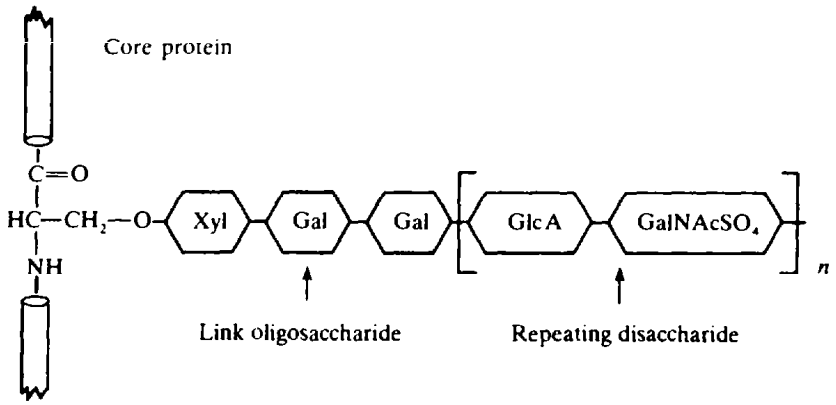


Fig. 5-21 Schematic representation of the linkage between the glycosaminoglycan chondroitin sulfate and a serine residue of the core protein.

during the elongation of the chain, sulfate groups are added, usually to a hydroxyl group of the amino sugar, to increase the negative charge of the molecules. In cartilage, the most abundant glycosaminoglycan is *chondroitin sulfate* ($M_r \approx 10{,}000$–$30{,}000$), and there may be over a hundred such chains linked to the core protein. *Sulfation* may occur either on the 4′ or 6′ hydroxyl group of the galactosamine.

EXAMPLE 5.14

As well as chondroitin sulfate, other glycosaminoglycans, such as *keratan sulfate*, with shorter chains ($M_r \approx 2{,}500$–$5{,}000$) are found in cartilage (Fig. 5-22). N-linked oligosaccharides similar to those found in glycoproteins are also present. Note that keratan sulfate in other tissues, e.g., intervertebral disk or cornea, has a longer chain ($M_r \approx 10{,}000$–$25{,}000$) and may have a different oligosaccharide link to the core protein.

There is still a further level of assembly, for, in the presence of hyaluronan, most cartilage proteoglycan molecules aggregate to form massive complexes, containing some 50 proteoglycan units, to give a total mass around 10^8 Da. The globular heads of the core proteins interact with segments of five disaccharide units along the length of the hyaluronan. This noncovalent interaction is stabilized by the binding of a *link protein* ($M_r \approx 50{,}000$) to both the hyaluronan and the core protein (Fig. 5-23).

The physical properties of proteoglycans are due almost entirely to the glycosaminoglycan components; the core proteins act largely as *spatial organizers*. The number and length of the glycosaminoglycan chains can vary considerably. The high density of negative charges causes electrostatic repulsion, so that the glycosaminoglycan chains are fully extended and separated to form a *bottlebrush*-like structure (see Fig. 5-24), which occupies a relatively large volume. The hydrophilic

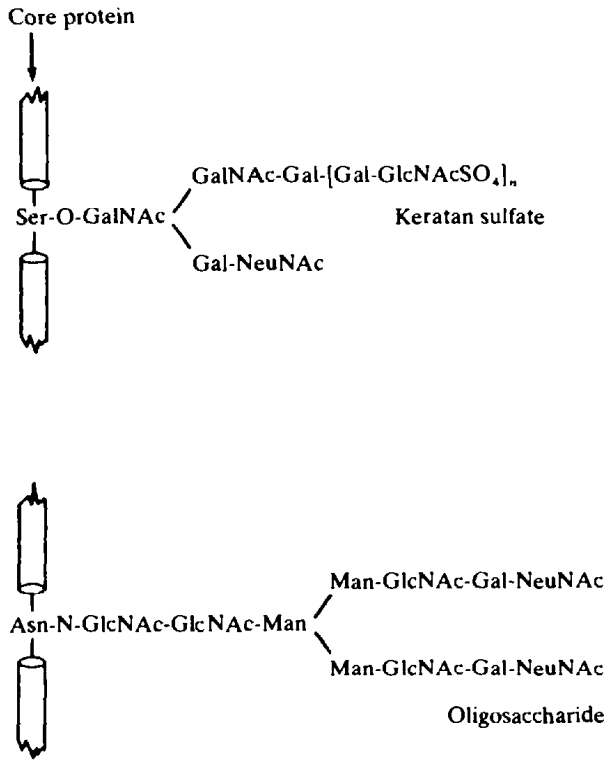


Fig. 5-22 Additional linked carbohydrates found in cartilage.

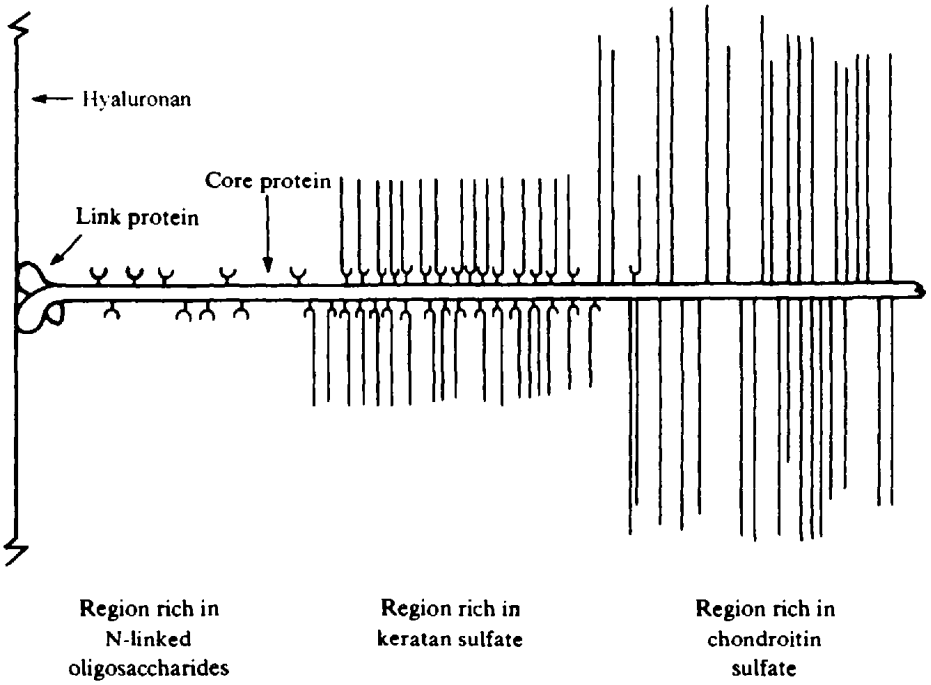


Fig. 5-23 A diagrammatic representation of part of a proteoglycan aggregate from cartilage.

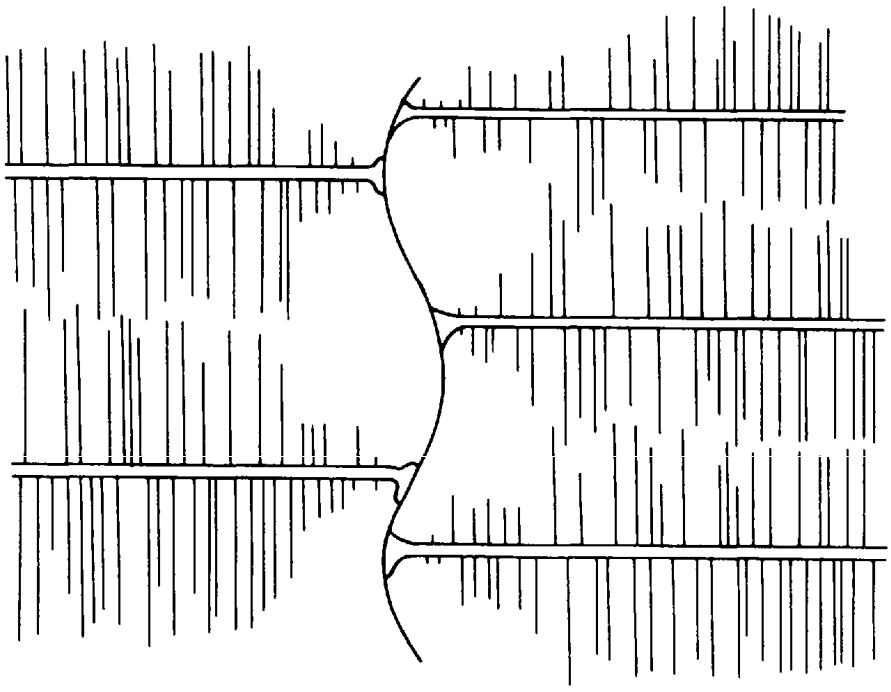


Fig. 5-24 Portion of a proteoglycan aggregate showing the bottlebrush-like structure.

groups of the glycosaminoglycans bind and immobilize large numbers of water molecules; the negative charges attract cations. These two effects create a swelling pressure or *turgor*.

EXAMPLE 5.15

Hyaluronan is the most studied of the glycosaminoglycans; it is also the largest. In aqueous solution, it forms a random coil and occupies a huge domain, whose volume is filled with immobilized water and to which small molecules and ions have access but large ones do not. In a 0.01% solution, the hyaluronan domains will occupy the total volume of the water; at higher concentrations, the hyaluronan molecules will interpenetrate neighboring domains, producing very viscous solutions that have lubricating properties, such as are needed in joints.

The space-filling character of glycosaminoglycans appears to be important in morphogenesis, particularly in the development of the skeleton. During these developmental processes, the presence of hyaluronan appears to facilitate the migration of cells. This effect is stopped by the removal of hyaluronan by hyaluronidase and by its replacement with aggregating proteoglycans.

Question: What are the functions of proteoglycans?

The large volume-to-mass ratio of proteoglycans, due to their ability to attract and retain large amounts of water, produces a swelling osmotic pressure, or turgor, in the extracellular matrix that resists compressive forces. This is demonstrated clearly in tissues such as joint cartilage and intervertebral disks. In *degenerative diseases* such as arthritis, proteoglycan is partially depleted and disaggregated, leading to changes in tissue resilience. The polyanionic bottlebrush structure of proteoglycans produces a sieving effect, so that the diffusion of macromolecules through connective tissue is restricted while small molecules, especially if anionic, may even show *enhanced rates* of diffusion. Other proteoglycans are much smaller and may contain only a few (1 to 10) glycosaminoglycan chains. Many are integral components of plasma membranes where they bind and regulate a variety of secreted proteins and can act as co-receptors for several growth factors.

Fibronectin

Of the several other protein components of the extracellular matrix, the best understood is *fibronectin*. This glycoprotein is a *heterodimer* composed of two very similar, but not identical, disulfide-bonded polypeptide chains ($M_r = 220{,}000$). It is found in several forms: as a bound complex on the surface of cells such as fibroblasts; as large aggregates in the extracellular space; and in a modified form as the so-called *cold-insoluble globulin* (i.e., it readily precipitates at 0°C) in plasma.

Fibronectin is a *multifunctional* protein consisting of a series of globular domains connected by flexible segments that are sensitive to proteolytic cleavage. These domains carry binding sites for components of the extracellular matrix and for cell surfaces (see Fig. 5-25). Fibronectin can thus form multiple interactions that are important in tissue homeostasis. For example, sulfated glycosaminoglycans increase the rate of binding of fibronectin to collagen, thus increasing the stability of the complex and allowing it to precipitate, forming large aggregates in the matrix.

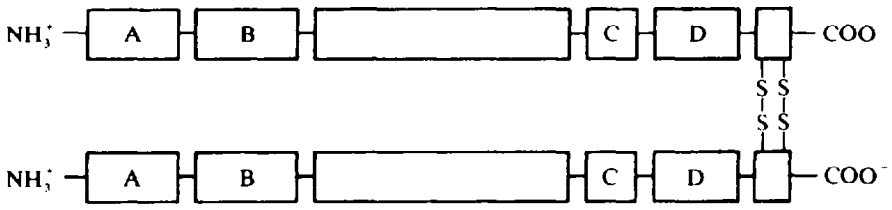


Fig. 5-25 Diagrammatic representation of the domains of fibronectin:
Domain A binds to heparin and fibrin.
Domain B binds to collagen.
Domain C binds to cell surfaces.
Domain D binds to heparin.

EXAMPLE 5.16

An important function of fibronectin is that of cell *adhesion*. Fibronectin forms a network that connects the cell to other components of the extracellular matrix, especially collagen, anchoring the cell in position. The connection to the cell is through a membrane protein, an *integrin*, that connects with the cytoskeleton. The minimal structure in fibronectin required for binding to an integrin is the sequence Arg-Gly-Asp (RGD) which is located at the apex of a surface loop of a *module* of the protein. *Transformed* cells produce less fibronectin and fail to adhere. The addition of fibronectin to cultures of such cells causes adherence, and the cells change to a more normal appearance. *Metastasizing* cancerous cells lack fibronectin or cause its proteolytic breakdown.

Structural studies have indicated that fibronectin is composed of at least three types of *modules*, internal amino acid sequence homologies, that have been found in several other protein sequences. Two other *attachment glycoproteins* have been described recently, and their properties are compared with fibronectin in Table 5.3.

5.6 THE CYTOSKELETON

The complex network of protein filaments found in the cytoplasm of eukaryotic cells is called the *cytoskeleton*. Some of the more important and better studied roles include: (1) providing a cell with its appropriate shape and reinforcing plasma membranes against *lysis* and *vesiculation*; (2) controlling changes in cell shape; (3) allowing cells to migrate (cell *motility*), e.g., during the movement of *macrophages* and *granulocytes* to sites of infection; and (4) transportation of large amounts of material

Table 5.3. Comparison of Attachment Proteins

Property	Fibronectin	Laminin	Chondronectin
Molecular weight	440,000	820,000	~170,000
Subunits	2 × 220,000	2 × 200,000 1 × 400,000	3 × ~55,000
Carbohydrate content	5–9%	12–15%	~8%
Shape	Extended V shape	Cross shape; 3 short arms, 1 long arm	Compact
Tissue distribution	Fibrous connective tissue	Basement membranes	Cartilage, vitreous body
Collagen-binding	Types I–V; best III and I	Type IV	Type II
Glycosaminoglycan-binding	Heparin, heparan sulfate	Heparan sulfate, heparin	Chondroitin sulfate, heparin
Cell-binding	Fibroblasts	Epithelial and endothelial cells	Chondrocytes

within, into and out of cells. These networks are therefore critical to cell *homeostasis*, and in sensing, responding to, and sending signals. For example, *endocytosis* at the plasma membrane allows certain nutrients and other material to be taken into a cell, packaged in small *vesicles*. *Signal transduction mechanisms* (Chap. 6), which control the response of a cell to external stimuli, often involve endocytosis. *Fast* and *slow transport* mechanisms inside cells distribute nutrients, cellular material and signalling molecules to appropriate areas of the cell. Material can be exported from cells in vesicles by *exocytosis*. A specialized form of exocytosis, *synapsis*, permits neurons to rapidly signal to each other.

Genetic errors in or damage to cytoskeletal components can play a central role in disease including many forms of hemolytic anemia and the *Duchenne* and *Becker muscular dystrophies*. Appropriate cell–cell contact (regulated through the cytoskeleton) appears to be critical in preventing cells from becoming cancerous and, indeed, many types of cancerous cell exhibit abnormal cytoskeletons.

A remarkable feature of cytoskeletal networks is that common elements can be incorporated into widely different structures in various parts of a cell. Networks can be dismantled in one part of a cell while other networks are being constructed from similar components in another. In fact, some networks undergo construction at one of their ends while simultaneously being dismantled at the other.

Microtubules, Intermediate and Thin Filaments

There are three main types of filament commonly found in the cytoskeletons of most cells. These are (1) *microtubules* (25 nm in diameter) composed principally of *tubulin*, (2) *thin filaments* (or *microfilaments*, 7 nm in diameter) which consist largely of polymerized *actin* and, (3) *intermediate filaments* (10 nm in diameter) which are manufactured from several classes of related proteins. Both tubulin and actin are globular proteins and are evolutionarily highly conserved. Their filaments tend to be much more dynamic than intermediate filaments (see Section 5.1).

All three types of filament are *polar* helical rods (i.e., they have chemically distinct "heads" and "tails"). In addition to the multiple contacts between subunits that serve to stabilize the filaments, additional multiple binding sites are available to bind *auxiliary* proteins. These auxiliary proteins can regulate the assembly and disassembly of the filaments, link filaments to one another and to other cellular components including membranes, and allow the filaments to participate in energy-dependent movement.

Actin Filaments

Actin ($M_r \approx 41{,}800$) is widely distributed in eukaryotic cells, often being the most abundant protein, and commonly making up about 10 percent of the total cell protein. The protein is highly conserved with only 17 of the 375 amino acids different between slime mold actin and rabbit muscle actin. Monomeric actin is usually referred to as globular or G-actin while the polymer, filamentous actin, is termed F-actin.

Actin contains a binding site for divalent cations (Ca^{2+} or Mg^{2+}) which must be occupied in order to prevent the protein from denaturing. It can also bind ATP, or ADP and inorganic phosphate (P_i). *In vitro*, actin can be maintained in the G-form by keeping the *ionic strength* very low. However, when the salt concentration is increased in the presence of ATP or ADP, polymerization to F-actin can occur.

Polymerization with ADP

The polymerization process in the presence of ADP is shown in Fig. 5-26. When the salt concentration is raised there is a characteristic *lag period*, usually of many minutes, before F-actin appears. The lag period results from the fact that sufficient *nuclei* (containing 2–4 monomers) must form before polymerization takes place. The nuclei are unstable, probably because they do not contain all of the subunit–subunit contacts that stabilize monomers in a filament (see Fig. 5-8). Once sufficient nuclei are available, polymerization proceeds rapidly from both ends of the nuclei. The filaments themselves can contain several hundred actin monomers, each with a molecule of ADP bound. The structure can be regarded as a right-handed double helix with 13.5 monomers per turn giving a helical repeat of 36 nm (see Fig. 5-8).

Actin will only polymerize if its concentration exceeds the *critical concentration* (C_c). Thus, at *chemical equilibrium*, a pool of G-actin remains whose concentration equals C_c. The equilibrium is *dynamic*; at the critical concentration the rate at which actin monomers bind to the ends of the filaments is equal to the rate at which they dissociate. If the system is diluted, however, there will be a net loss of monomers from the filaments and this *depolymerization* will continue until the concentration of G-actin is restored to the critical concentration.

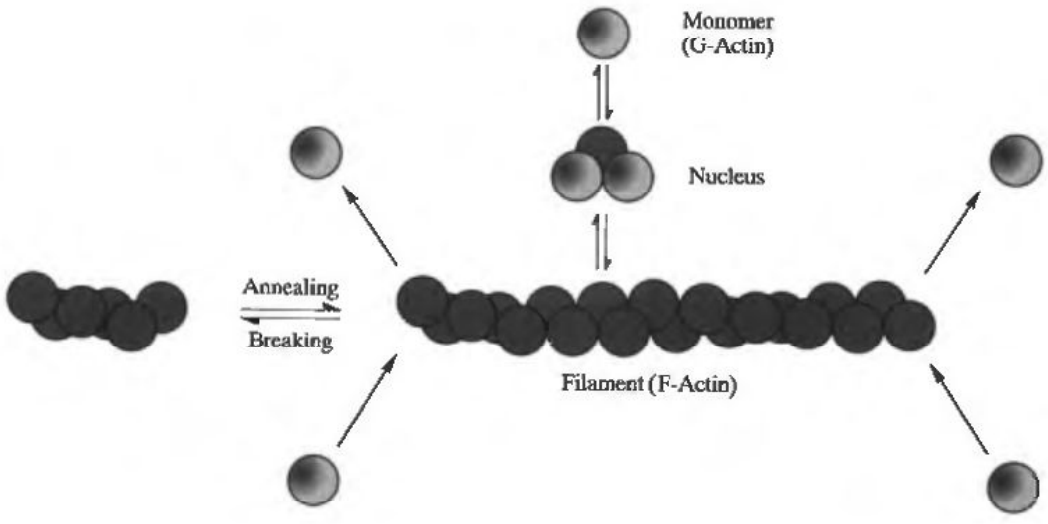


Fig. 5-26 Assembly and disassembly of actin filaments in the presence of ADP.

Question: How can the *polarity* of an actin filament be determined?

The protein *myosin* (Fig. 5-32) interacts specifically with each actin molecule in a filament and, as a result, the filament becomes "decorated" with a pattern of arrow-heads all pointing the same way. Because of this pattern, one end of the filament is known as the *pointed end* while the other is called the *barbed end*.

Polymerization with ATP

Polymerization in the presence of ATP, as opposed to ADP, is more complex but physiologically relevant. The main differences lie in the properties of F-actin (Fig. 5-27). The two ends of the filament now become *kinetically distinct*. When polymerization is initiated, monomers with ATP bound to them add much more rapidly to the barbed end than to the pointed end.

The ATP bound to each actin monomer is hydrolyzed to ADP + P_i soon after the monomer binds to the filament. Thus, most of a growing filament contains actin with bound ADP, while the ends contain small "caps" of ATP-actin. This *irreversible* hydrolysis of ATP means that the system can never reach chemical equilibrium while the ATP is present in solution. Instead, a *steady state* is achieved with a *different* value of C_c applying to each end; the value of C_c is much lower at the barbed end than at the pointed end. The concentration of G-actin at steady state lies *between* the C_c values for each end.

Because of these differences in C_c, and the energy provided by the hydrolysis of ATP, the filaments at steady state can *treadmill*. There is a net loss of monomers (mostly ADP-actin) from the pointed (or *minus*) ends since the value of C_c is higher than the concentration of G-actin. There is a net gain of ATP-actin at the barbed (or *plus*) ends since the value of C_c is lower than the concentration of G-actin.

ADP-actin released from the filaments can exchange ADP for ATP in solution and rebind to the filaments. Due to a difference in conformation, monomers in filaments cannot exchange their ADP for ATP.

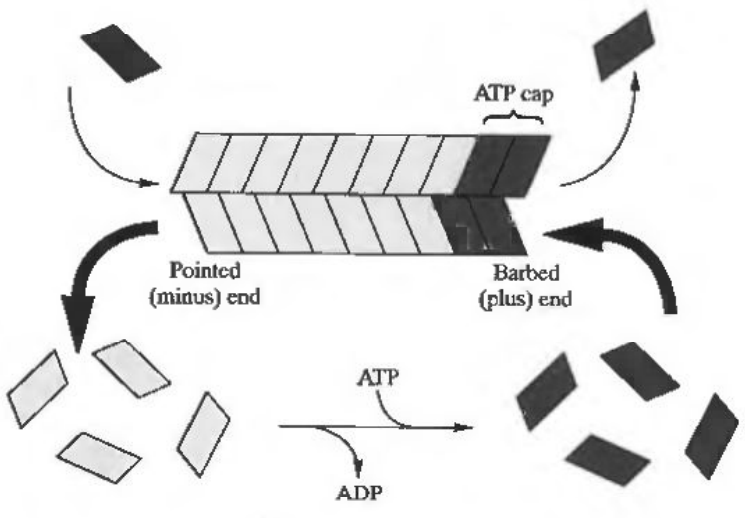


Fig. 5-27 Steady-state properties and treadmilling of actin filaments in the presence of ATP.

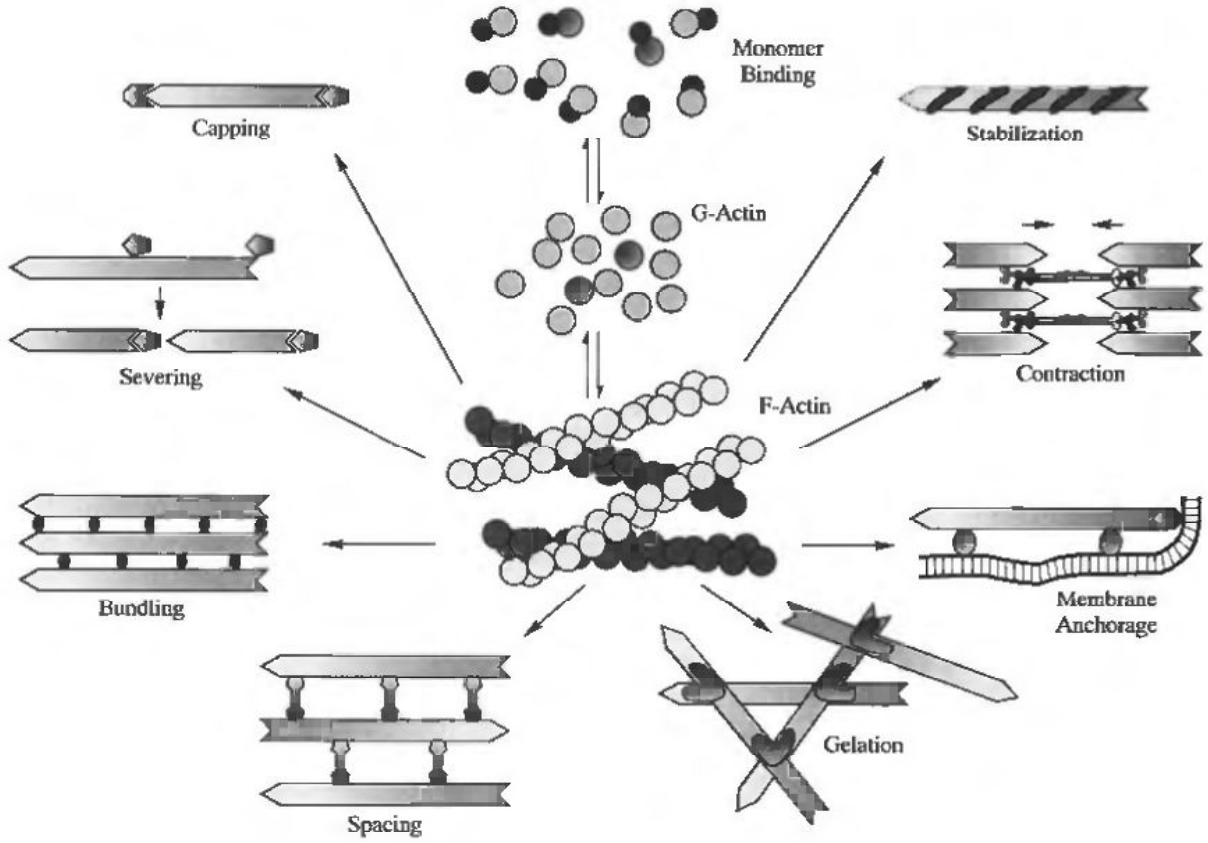


Fig. 5-28 The effect of different types of actin-binding protein on the organization of actin.

Actin Binding Proteins and the Control of Actin Polymerization

Actin filaments are found throughout the cytoplasm of most cells, although, in general, a large percentage of the actin is concentrated at the *cell cortex*, the region just beneath the plasma membrane. The filaments are arranged with respect to one another in a great variety of different ways through the actions of *actin binding proteins*. These proteins, their regulators, and a range of other mechanisms, serve to control actin polymerization and the stability and dynamics of actin-based networks.

Actin binding proteins can be divided into several classes, depending on where they bind to the filament and the roles that they play; these classes are defined and examples given above (Fig. 5-28).

Monomer Binding

Several proteins are known to bind specifically to actin monomers. In most cases these monomer-binding proteins, as expected, inhibit polymerization, promote the depolymerization of preformed filaments, and raise the critical concentration, C_c.

EXAMPLE 5.17

In the fibroblast, and other cells, the concentration of monomeric actin can be as high as 200 μM and represent up to 50% of the total actin in the cell. This contrasts sharply with the critical concentration of ATP-actin *in vitro* which is typically 1 μM or less. *Thymosin* (M_r = 5,000) is one protein that is responsible for *sequestering* monomeric actin in these cells. A large pool of monomeric actin provides the opportunity to rapidly assemble an actin-based cytoskeletal network while leaving preexisting actin-based networks intact.

Capping Proteins

Certain proteins can specifically bind to either the barbed or pointed ends of actin filaments. Capping the filaments in this way prevents, or retards, any further elongation or depolymerization of monomers from that end.

EXAMPLE 5.18

The best studied example of a barbed-end capping protein is *gelsolin* (M_r = 90,000). Like many actin-binding proteins, gelsolin's affinity for actin filaments is regulated by changes in the concentrations of Ca^{2+} which occur when cells are stimulated in particular ways. Approximately 10^{-6} *M* Ca^{2+} stimulates barbed-end capping and promotes the depolymerization of actin, since the observed critical concentration of the system now increases towards that of the pointed end alone.

Severing Proteins

Several proteins are now known to sever actin filaments and rapidly convert cross-linked F-actin *gels* into less viscous *sols*. These gel-to-sol transitions are usually Ca^{2+}-regulated and are observed in the cortex of many cells.

EXAMPLE 5.19

Gelsolin can sever actin filaments in addition to binding their barbed ends. The severing activity is also Ca^{2+} dependent and involves substantial bending of the actin filaments. Following severing, gelsolin can prevent *reannealing* of the filament by binding to the newly created barbed end.

Side-Binding Proteins

Proteins which bind exclusively to the sides of the actin filament (as opposed to capping the ends) tend to promote the formation of filaments and lower the critical concentration. In some cases, the mechanical stability and stiffness of the filament are also enhanced.

EXAMPLE 5.20

Tropomyosin is a *coiled-coil* dimer consisting of identical 35 kDa subunits. It binds to the grooves of the actin filament spanning approximately seven actin subunits and serves to stiffen and straighten the filaments. The protein binds F-actin cooperatively, meaning that once one molecule binds, the entire length of the filament is likely to bind tropomyosin in its grooves.

The function of tropomyosin is best understood in the actomyosin contractile apparatus of striated muscle where it serves as part of the triggering mechanism for muscle contraction (see *Force-Generating Proteins* below). In resting muscle, tropomyosin is held out of the groove of the actin filament by the *troponin complex*, which consists of three polypeptides. This blocks the interaction of actin with the heads of the myosin molecules (Fig. 5-29). When calcium is released into the cytoplasm from intracellular stores following a nerve stimulus, it can bind to *troponin C*, altering the conformation of the troponin complex. The tropomyosin can then move towards the groove exposing the myosin binding site on actin and allowing the binding of myosin to actin.

Cross-Linking Proteins

Cross-linking proteins also bind to the sides of actin filaments but can simultaneously bind more than one filament. There are several important subclasses of these proteins which arise from the different modes of cross-linking. *Bundling* proteins, such as *villin* and *α-actinin*, cross-link actin filaments into regular bundles with narrow gaps (<40 nm) between the filaments. *Spacer* proteins, such as the flexible, highly elongated *spectrin* tetramer, give rise to more loosely packed filament aggregates with gaps of up to 200 nm between filaments. *Gelation cross-linkers*, such as the filamin dimer, bind two filaments together at the point at which they overlap, regardless of the angle of the filaments with respect to one another.

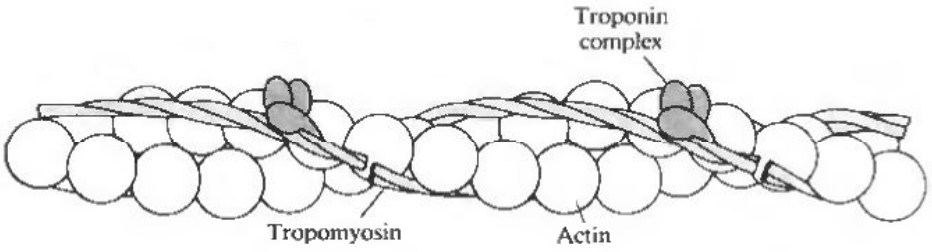


Fig. 5-29 Model of the interaction between F-actin, tropomyosin, and the troponin complex in resting muscle.

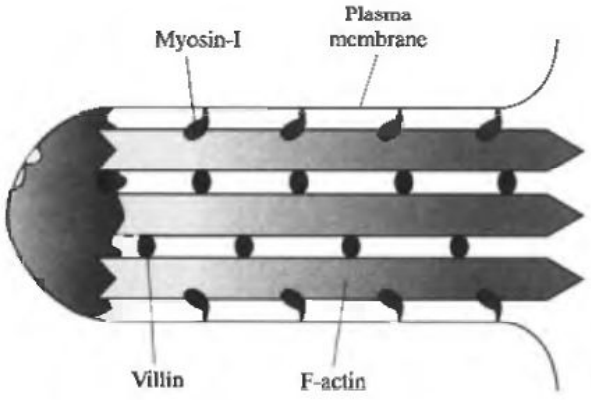


Fig. 5-30 Bundling and membrane anchorage of F-actin in a microvillus.

EXAMPLE 5.21

Villin is an example of a *bundling* protein. Villin is found in the microvilli of, for example, intestinal *brush border* cells (Fig. 5-30). The microvilli greatly increase the surface area of the cells, which is essential for effective absorption to take place. Each microvillus extends about 2 μm into the lumen of the gut and is supported by 20 or so actin filaments tightly bundled by villin (and other proteins) at regular intervals. In a feature common to many actin-based networks, all the filaments in the bundle are oriented with their barbed ends in the same direction, in this case toward the tip of the microvillus where they terminate. Cross-linking of the actin filaments to the plasma membrane occurs via a second protein from the myosin-I family (a relative of the well-known contractile protein myosin-II). This protein binds its "head" domain to the sides of the filaments and embeds its "tail" domain into the membrane.

EXAMPLE 5.22

The cytoskeleton of mature mammalian erythrocytes is as unusual as the highly differentiated cells themselves. During maturation of the erythrocyte, the cytoskeleton becomes restricted to a two-dimensional network lining the cytoplasmic face of the membrane (Fig. 5-31). This flexible and durable *membrane skeleton* plays a major role in preventing the cell from breaking up during its arduous passage through the circulation while allowing it to be flexible enough to squeeze through very narrow capillaries.

The main component is the spectrin heterotetramer (two α and two β chains), a very large (1,052 kDa), highly elongated, flexible molecule which can cross-link actin using the binding sites at its ends. *Protein 4.1* greatly strengthens the interaction between spectrin and actin. Unlike other cells, the actin in erythrocytes forms *protofilaments* consisting of only approximately 15 subunits. On average, six spectrins attach to each protofilament, forming a largely hexagonal interlocking lattice. This lattice is attached to the membrane primarily through *ankyrin* which binds to both the β chain of spectrin and the membrane-inserted *anion channel*.

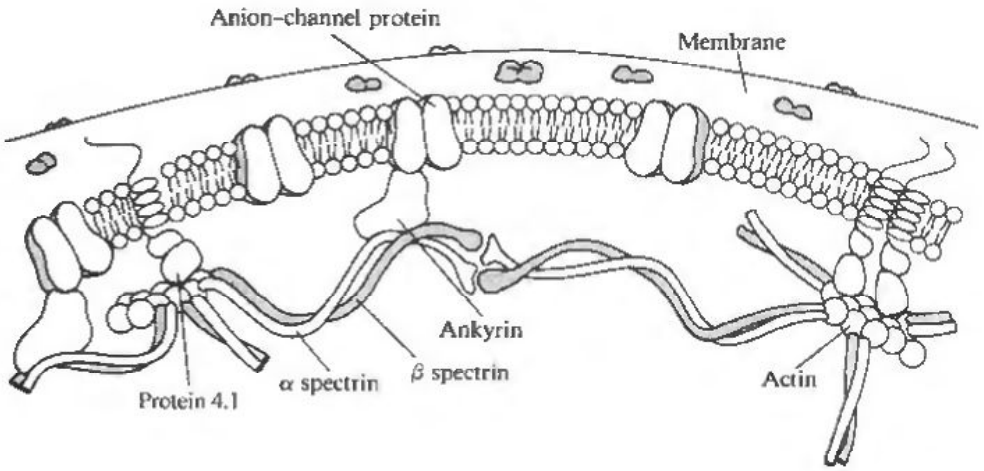


Fig. 5-31 The membrane skeleton of a mature mammalian erythrocyte.

Force-Generating Proteins

Myosin-II, the force-generating protein of the contractile apparatus of muscle and other cells, was the first actin-binding protein described and characterized. The myosin protomer ($M_r \approx 520{,}000$) consists of two identical *heavy chains* ($M_r = 220{,}000$) and two pairs of light chains ($M_r = 16$–$27{,}000$ depending on the source). Each heavy chain contains a globular head linked by a hinge region to a long α-helical tail. The tails from the two heavy chains form a coiled-coil 150 nm long (Fig. 5-32). The light chains bind near the base of each head.

The myosin protomers can self-associate to form a *thick filament* by the assembly of some 400 myosin tails in a staggered side-by-side packing with the myosin heads projecting at regular intervals in a helical array (Fig. 5-32). The thick filament is *bipolar* with a 150 nm bare zone in the center where two oppositely oriented sets of myosin tails come together. The center of this region is called the *M line*. This thick filament forms part of a *myofibril*.

The myofibril also contains *thin filaments* of F-actin with tropomyosin dimers bound in the grooves and about one troponin complex per tropomyosin dimer. The thin filaments are also arranged in a bipolar fashion. All the pointed ends point away from each side of the *Z-disk* which contains a number of proteins including α-actinin. α-Actinin serves to bundle the actin filaments very close to their barbed ends and anchor them very firmly in the disk.

The thick filaments and the thin filaments of a myofibril partly interdigitate with six thin filaments surrounding each thick filament. A single contractile unit, the *sarcomere*, is depicted in Fig. 5-32 and is about 1.5 μm in diameter and 2.2 μm long. A linear arrangement of sarcomeres forms the myofibril, which, along with other myofibrils packed in parallel, can run the length of the muscle cell (up to several centimeters).

Upon contraction, the thick and thin filaments in each sarcomere of each myofibril coordinately slide past each other without changing length. The sarcomeres shorten and the interdigitation between thick and thin filaments increases. This is the *sliding filament model* of muscle contraction.

Question: What is the mechanism by which myosin generates mechanical force between thick and thin filaments?

Mechanical force is generated by the cyclic interaction of myosin heads with actin and the energy supplied from the hydrolysis of ATP (Fig. 5-33). Each myosin head can bind a single molecule of ATP. Myosin is an *ATPase*. The ATPase activity is *constitutive* but can be increased up to 200-fold in the presence of actin.

The cycle begins with the myosin head hydrolyzing the bound ATP to ADP and P_i. In this form,

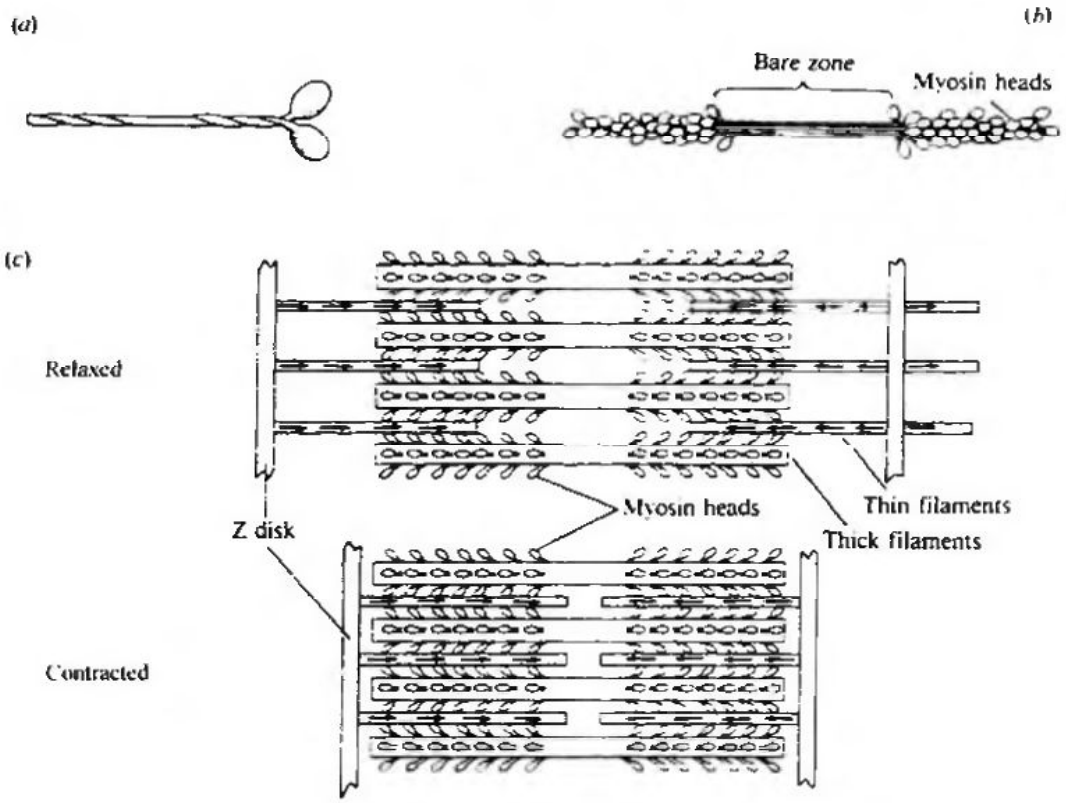


Fig. 5-32 Components of a skeletal muscle myofibril: (*a*) myosin, (*b*) thick filament, and (*c*) the contraction of a sarcomere. The polarities of the thin filaments are indicated by the arrows.

the head binds to the side of an actin filament which stimulates the release of products, ADP and P_i, and causes a change in conformation around the myosin head. The head pulls back, dragging the actin filament with it. This is the *power stroke*. The head then detaches from the actin and binds a new molecule of ATP to begin the cycle over again. Each cycle can be as short as 0.2 s during rapid contraction. Note the importance of the two hinges which permit flexibility of movement.

Question: How is the generation of mechanical force initiated and regulated?

An increase in Ca^{2+} (e.g., from $10^{-8}\,M$ to $10^{-5}\,M$) acts as the trigger. In striated muscle, Ca^{2+} is released from the *endoplasmic reticulum* into the cytoplasm following stimulation of the muscle cell via its attached motor nerve. The Ca^{2+} interacts with the troponin complex, causing a movement of tropomyosin to expose the myosin binding sites on the thin filaments (see Example 5.21). In *smooth* muscle, the released Ca^{2+} indirectly activates myosin light chain kinase which phosphorylates the light chains of myosin. Hence, the control is at the level of the thick filament. In some nonmuscle cells, the control by Ca^{2+} is at the level of the assembly of myosin into filaments.

In all cases, a fall in Ca^{2+} concentration (e.g., by being pumped back into intracellular stores) reverses the process. It should be noted that a continuing supply of ATP is required for the contraction process. When ATP is completely depleted, *rigor* sets in.

Question: What regulates thin filament length and the position of the thick filaments in the sarcomere?

Thin filaments in the sarcomere are of remarkably uniform length. In recent years, two proteins, *nebulin* and *tropomodulin*, have been discovered in the sarcomere. Nebulin is a huge protein consisting almost entirely of multiple 35-residue *actin-binding motifs*. Nebulin's length is that of the thin filaments of the sarcomere and therefore it may serve to regulate thin filament length via the

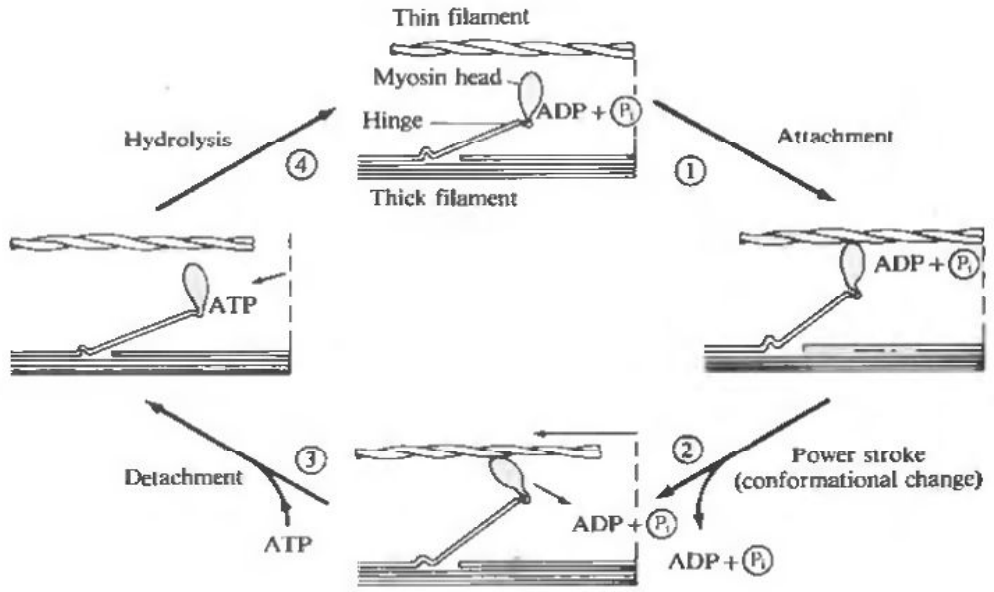


Fig. 5-33 Representation of the four steps in the contractile cycle of muscle showing the association and dissociation of myosin and actin driven by the hydrolysis of ATP.

vernier principle (Fig. 5-4). Tropomodulin participates in this process by capping the pointed ends of the filaments, preventing them from growing further, and is thus an example of a pointed-end capping protein.

Another recently discovered protein, *titin*, is also extremely large and highly elongated. Titin molecules extend from the Z-disk to the thick filaments and may serve to center these filaments in the sarcomere.

Microtubules

Microtubules are hollow cylinders with an internal diameter of 14 nm and an external diameter of 25 nm. A complete cylinder contains 13 parallel *protofilaments* each of which is a head-to-tail polymer of heterodimers forming a rod (Fig. 5-34). The heterodimers contain an α- and a *β-tubulin*

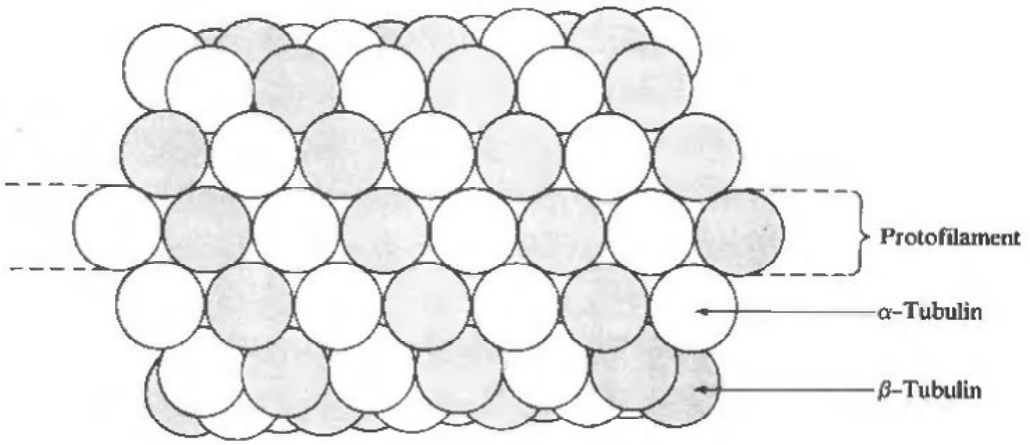


Fig. 5-34 A short segment of microtubule showing the tubulin dimers aligned into protofilaments.

subunit tightly bound together. These two subunits (each $M_r \approx 50{,}000$) are synthesized from separate genes but their amino-acid sequences are highly homologous.

Polymerization, Treadmilling and Dynamic Instability

The polymerization of tubulin has many similarities to that of actin. The protomer is the tubulin heterodimer. Each subunit in the heterodimer can bind a nucleotide; in this case it is GTP. There is a lag period associated with nucleation. Once sufficient nuclei are available, "explosive" polymerization can proceed from both ends of the nuclei. Calcium ions and low temperature (4°C) inhibit polymerization, while magnesium ions and high temperature (37°C) stimulate it. As with actin, the microtubule is polar, and polymerization proceeds about three times more quickly from the *plus end* than the *minus end*. Polymerization is accompanied by hydrolysis of GTP, on the β-tubulin only, at or soon after the addition of the α/β protomer to the filament. *GTP caps* can therefore exist at both ends of the filament with the cap likely to be larger, on average, at the plus end than at the minus end.

At steady state, a pool of free heterodimers (representing the critical concentration) coexists with filaments. For these filaments, the critical concentration is lower at the plus end than at the minus end. Thus, treadmilling can occur with a net loss of protomers from the minus end and net addition at the plus end. However, for microtubules, treadmilling is generally overshadowed by a related and much more dramatic phenomenon known as *dynamic instability*. If a filament loses its GTP cap at either end, depolymerization at that end occurs extremely rapidly. Thus, even at steady state, some filaments can be rapidly shortening at one or both ends while other filaments are rapidly growing at one or both ends. Dynamic instability also occurs *in vivo* and is an essential factor in the proper functioning of many microtubule-based networks.

Microtubule-Associated Proteins (MAPs)

MAPs can be classified in a similar manner to actin-binding proteins. Like their actin-binding counterparts, MAPs can nucleate, cap, stabilize and cross-link microtubules. Several *motor proteins* have been identified in the last few years which are involved in force generation in a number of microtubule-based networks. As with actin-binding proteins, MAPs can be targeted, on a constitutive or transient basis, to certain regions of a cell. This allows specialized microtubule-based networks to be constructed and perform their tasks in specific locations. The properties and functions of some of these MAPs will be explored in the examples given below.

Question: How are microtubules distributed inside cells?

During the *interphase* of almost all animal cells, most of the microtubules radiate from the *centrosome*, a darkly staining region near the nucleus. The centrosome contains a pair of *centrioles*, which themselves are small cylinder-shaped structures containing microtubules and other proteins [see Fig. 5-36(*a*)]. Surrounding the centriole pair is the *pericentriolar material* which serves somehow to nucleate the formation of new microtubules. For this reason the centrosome is known as one type of *microtubule organizing center*, or *MTOC*. The microtubules radiate with their plus ends toward the periphery of the cell and their negative ends embedded, and stabilized in the centrosome. The microtubules remain highly dynamic *in vivo* with a half-life of the order of a few minutes. Some microtubules might be stabilized by, for example, having their plus ends capped, thus preventing depolymerization.

EXAMPLE 5.23

The interphase organization of microtubules serves many roles, among the most important of which is the rapid transport of organelles and materials packaged in vesicles to various parts of a cell. This process was first observed directly in the giant axons of squid and was therefore named *fast axonal transport*. In highly elongated

cells, such as neurons, where simple diffusion of organelles and even proteins would be unacceptably slow, this rapid transport mechanism serves to deliver essential material to distal parts of the cell and to retrieve other material. *Anterograde* movement of small vesicles in the axon (i.e., away from the cell body and toward the plus end of the axonal microtubules) can be as fast as 5 μm per second. *Retrograde* transport can be as much as half this rate.

Fast transport is now known to occur in almost all animal cells. The microtubules act as "superhighways" upon which large numbers of vesicles, and other organelles such as mitochondria, move.

Question: What is the mechanism of fast transport?

Fast transport in cells requires vehicles (the organelles or vesicles), motors and fuel. The motors can be divided into two families, the *cytoplasmic dyneins* and the *kinesins* (Fig. 5-35). Each uses ATP as a fuel, the energy being provided by its hydrolysis to ADP and P_i. The motors can sense the polarity of the microtubules; dyneins drive toward the minus end, while most kinesins drive toward the plus end. Movement is generated by cyclic hydrolysis of ATP, conformational changes, and the reversible attachment to microtubules.

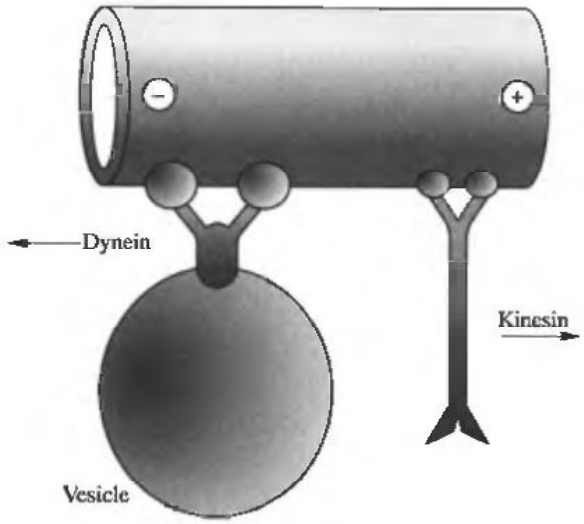


Fig. 5-35 Dynein and kinesin motors and their interaction with vesicles and microtubules.

EXAMPLE 5.24

Cilia and *flagella* are stable microtubule-based structures which project from the plasma membranes of particular eukaryotic cells. The energy-dependent oscillations of these structures can drive material over the surface of a cell or propel the cell along. For example, the whip-like motions of cilia on the cells at the head of the fallopian tube draw newly released ova from the ovaries into and along the oviduct. The snake-like movements of the flagellum on a sperm provide these cells with movement.

Although the movements of cilia and flagella are somewhat different, the microtubule-based *axoneme* is common to them both. The axoneme is cylindrical (~200 nm in diameter) and built from the "9 + 2" arrangement of microtubules [Fig. 5-36(*b*)]. Note that the *two central* microtubules are complete while each of the nine *doublets of outer* microtubules contains one complete (A tubule) and one incomplete (B tubule) fused together.

A number of auxiliary proteins are associated with the axoneme at regular intervals. Every 24 nm, *ciliary dynein* (a three-headed isoform of cytoplasmic dynein) extends pairs of arms from the A tubule of each doublet toward the B tubule of the adjacent doublet. At intervals of 86 nm, *nexin* links adjacent doublets. Every 29 nm, *radial spokes* project inward from the outer doublets toward the *inner sheath* which itself is connected to the central pair of microtubules.

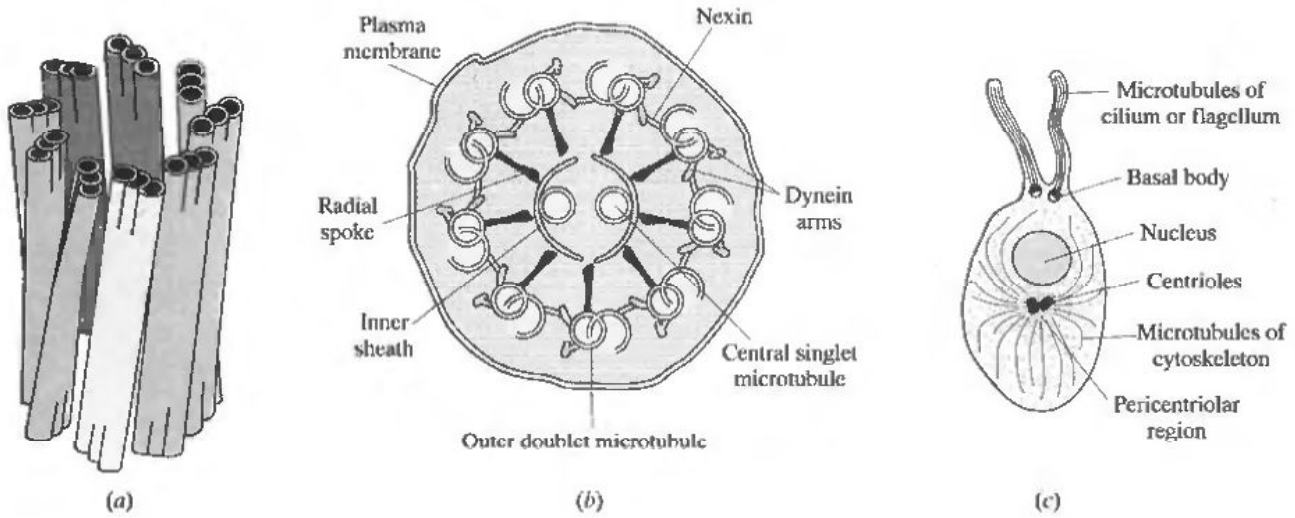


Fig. 5-36 Representations of centrioles, basal bodies and axonemes within a cilium. (*a*) Centriole/basal body showing the nine sets of triplet microtubules. (*b*) Cross-section of a cilium/flagellum showing the "9 + 2" array of microtubules and accessory proteins of the axoneme. (*c*) View of a cell indicating the organizing role of centrioles and basal bodies.

Without the cross-links, the dynein motors would simply cause the microtubules to slide past one another until they were no longer in contact. The cross-links limit sliding, and play a role in the poorly understood mechanism which promotes bending and cyclical beating.

Question: What organizes the structure of the axoneme?

An axoneme grows from a preexisting structure called the *basal body* which appears to be similar in construction to a centriole in the centromere [Fig. 5-36(*a*)]. The basal body consists of nine triplet sets of microtubules cross-linked by proteins at regular intervals. There are no central microtubules, but protein "spokes" radiating inward to a central "hub" can be seen in some cross-sections.

Each axoneme doublet grows from two of the tubules in each triplet of the basal body. It is not known how the growth of the two central microtubules of the axoneme is controlled, or how the carefully regulated length of the cilia and flagella is monitored.

EXAMPLE 5.25

The *mitotic phase* is the culmination of the *cell cycle* in which two identical daughter cells are produced each with a full complement of genetic material. The process is extremely complex and dynamic and lasts little more than an hour in a "typical" cell. During this time, the microtubule-based *mitotic spindle* (Fig. 5-37) must be constructed and the chromosomes condense (*prophase*). The *nuclear envelope* then disintegrates and the chromosomes are "captured" by microtubules in the spindle (*prometaphase*). By *metaphase*, the spindle has moved and aligned all of the chromosomes at the *metaphase plate*, a plane midway between the *spindle poles*. *Anaphase* begins when the connection between the *sister chromatids* of each chromosome suddenly breaks and the sister chromatids are moved swiftly toward opposite spindle poles. The poles also move apart. Once at the poles, the chromatids decondense and each set is enclosed by new nuclear envelopes (*telophase*). Two new cells are formed by *cytokinesis*, which begins in late anaphase, and involves a "pinching in" or cleavage of the plasma membrane at a point midway between the spindle poles.

Question: What is the source of the microtubules in the mitotic spindle?

The microtubule network of the interphase cell is depolymerized and reformed to make the mitotic spindle. This requires two MTOCs, the first being the original centromere with its pair of centrioles.

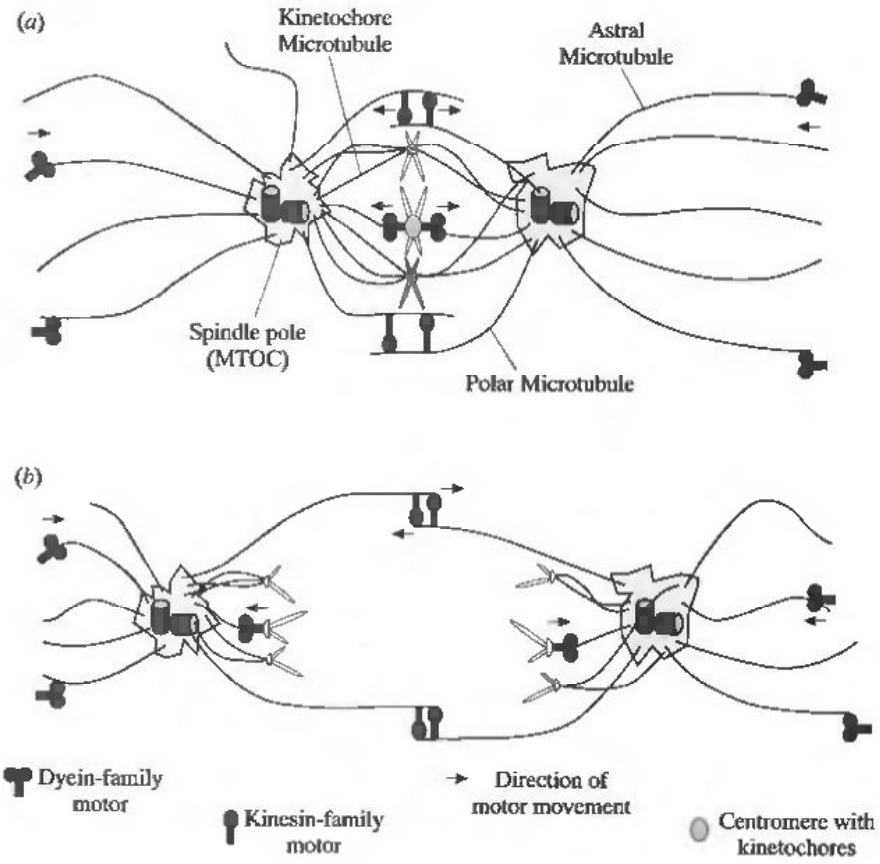


Fig. 5-37 Diagram showing (*a*) the mitotic spindle at metaphase with chromosomes aligned at the metaphase plate and (*b*) separation of the poles and sister chromatids during anaphase.

The second arises from replication of the centrioles during the S and G_2 phases of interphase. During prophase, the two pairs of centrioles move apart from one another and the pericentriolar material surrounding them both can nucleate and stabilize microtubules. These two MTOCs become the spindle poles.

Three types of microtubule can readily be defined in the mitotic spindle. *Polar* microtubules overlap (and probably interact) between the poles and are involved in pushing the poles apart in anaphase. *Astral* microtubules radiate in all directions and also help separate the poles. *Kinetochore* microtubules attach themselves to specialized protein structures (kinetochores) located on each side of the *centromere* of each chromosome. These microtubules are involved in moving the chromosomes to the metaphase plate and in separating sister chromatids at anaphase. The microtubules in the spindle are very dynamic and have a half-life of only a few seconds. This appears to be especially important in the capture of chromosomes by the kinetochore microtubules. Microtubules that miss the target kinetochores are quickly lost because their dynamic instability soon leads to depolymerization. The new microtubules that form may hit the target and be partially stabilized through plus-end capping.

Question: How are forces generated to move chromosomes and separate the poles?

The spindle contains a number of dynein- and kinesin-family motors (Fig. 5-37). These, together with the dynamic instability and treadmilling of the microtubules, help control the movement and relative positions of the chromosomes and the poles. The spindle is under constant *tension* resulting from a range of forces.

During metaphase, when all of the chromosomes are lined up at the metaphase plate, all of these forces are in balance. Dynein motors in the kinetochores attempt to drag the chromosomes back to the poles (i.e., in the minus direction) along the kinetochore microtubules. Since each chromosome contains two kinetochores, each facing a different pole, the pull to each pole cancels out [see arrows in Fig. 5-37(*a*)]. Simultaneously, at least two forces are acting to separate the poles. Where the polar microtubules overlap, kinesin motors are present binding their tails to the sides of the microtubules and using the motors in their heads to drive toward the plus-ends of others, thereby exerting a push on the poles. In astral microtubules directed away from the center of the spindle, it appears that dynein motors (perhaps attached to the cell cortex) exert a pull on the poles.

During anaphase, when the sister chromatids split at the centromere, this balance is lost [Fig. 5-37(*b*)]. The dynein motors in the kinetochores drive the sister chromatids to the poles. The speed at which this happens may be regulated by the controlled depolymerization of the kinetochore microtubules. The poles can now move apart and the overlap between polar microtubules shortens. It would be completely lost but for rapid polymerization at the plus ends of these filaments.

Question: How does cytokinesis occur?

Cytokinesis begins some time in anaphase. A *cleavage furrow* begins to form in the plasma membrane in the same plane as the metaphase plate. It is not known how the furrow is positioned. The furrow is created and deepens due to the actions of the *contractile ring* which is an actomyosin network assembled specifically for this purpose and dispersed afterward. Thus, the stability of this network is quite different from its analog in muscle. The dynamics of these networks are also different (see Prob. 5.16).

Intermediate Filaments

Intermediate filaments form a family of tough, insoluble protein fibers with diameters of about 10 nm, found in most eukaryotic cells. The protein subunits (Fig. 5-38) vary greatly in size (M_r = 40,000 to 200,000), but they all appear to have a common coiled-coil, α-helical region in the center of each subunit. The rodlike region is of constant length, with considerable sequence homology, 30–70 percent of the amino acids being identical between the various types; it also appears to be the basis for the filamentous structure, about 25 subunits interlocking side by side to form the cross-section of one filament. The terminal globular regions are the variable segments, which differentiate the various types of intermediate filaments and which project from the filament into the cytoplasm to play an as yet ill-defined role in the cell.

There are five classes (see Table 5.4) of intermediate filaments, which can be differentiated by biochemical, immunological, and molecular biological techniques. In general, one type of cell contains only one class of intermediate filament. All the classes are believed to play a structural role within the cell by resisting mechanical stress.

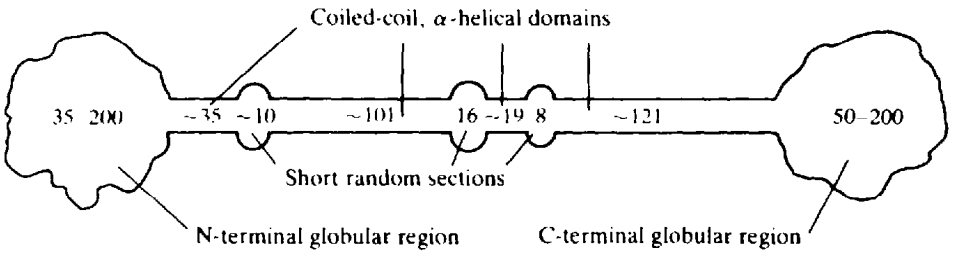


Fig. 5-38 A diagrammatic representation of the subunit of an intermediate filament. (The figures represent the number of amino acid residues in each structural domain.)

Table 5.4. The Classes of Intermediate Filaments in Vertebrate Cells

Intermediate Filaments	Protein Components	Cell Type
Cytokeratins	Various: M_r = 40,000 to 70,000	Epithelium
Neurofilaments	Three major: M_r = 68,000, 145,000, and 220,000	Neurons
Vimentin	M_r = 55,000	Mesenchyme, e.g., fibroblasts
Desmin	M_r = 53,000	Muscle
Glial filaments	M_r = 50,000	Some glial cells, e.g., astrocytes
Nuclear lamins	Three major: M_r = 65,000 to 70,000	Nuclear lamina of eukaryotic cells

EXAMPLE 5.26

The most-studied class of intermediate filaments, and the most diverse is the *cytokeratins*, also called the *prekeratins*, which are cross-linked by disulfide bonds. On reduction, several different polypeptide chains are obtained from epithelial cells, with a spectrum of chains that varies from tissue to tissue. Some 19 different cytokeratin polypeptides have been characterized from human tissues, and each has been shown to be the product of a distinct individual gene and not formed by differential processing. There are two groups, the *acidic keratins* and the *neutral-basic keratins*, which are believed to interact in a 1:1 ratio to form the keratin filaments. These play an important role in epithelial cells by forming a continuous network of fibers across a whole epithelial sheet, giving it *tensile strength*. The cells are fastened by rivetlike connections, called *spot desmosomes*, which act as anchoring points for the keratin filaments. Keratin fibers are abundant in the tough outer protective covering of higher animals, and they become increasingly cross-linked as the cell matures and finally dies, producing the keratinized outer skin layer, nails, and hair.

The intermediate filaments are generally more durable than F-actin and microtubules, and the pool of unpolymerized subunits is very small. However, their assembly and disassembly can be regulated in various ways, including proteolysis and phosphorylation.

EXAMPLE 5.27

The inner nuclear membrane of the nucleus in mammalian cells is supported by a network of intermediate filaments called the *nuclear lamina* which is comprised of *lamins*. In late prophase of mitosis, the nuclear membrane fragments into vesicles triggered largely by phosphorylation of the lamins at various serine residues. This serves to disassemble the nuclear lamina. At the completion of mitosis, dephosphorylation of the lamins allows the vesicles in each of the new daughter cells to reform nuclear membranes surrounding the chromosomes.

Solved Problems

ASSEMBLY OF SUPRAMOLECULAR STRUCTURES

5.1. A multienzyme complex has three different catalytic activities with eight sites for each activity. Compare the frequencies of defective complexes produced in the following two situations: (*a*) The complex is synthesized in one step as one long polypeptide chain containing 8,000 amino acid residues. (*b*) The complex is constructed in three steps. First, 24 polypeptides are

synthesized: 8×200, 8×300, and 8×500 amino acid residues. Next, trimers consisting of one of each chain type are formed. Last, these eight trimers are assembled to form the complex. (Assume in both cases that the error frequency is 10^{-5} for each operation and that a single mistake will cause complete rejection.)

SOLUTION

(*a*) The probability of incorporating an incorrect amino acid is 1 in 10^5. Therefore, the frequency of defective complexes would be $8 \times 10^3 \times 10^{-5}$, i.e., eight of every 100 complexes are defective.

(*b*) Because defective polypeptides will be rejected as being unable to form trimers and any faulty trimers will not assemble further, the frequency of defective complexes will be related only to the seven steps required for the final assembly of the trimers to form the complex, and this would be equal to 7×10^{-5}. Thus, the three-step process produces about 1,000 times fewer defective complexes than the single-step process and, incidentally, requires one-eighth as much genetic information.

PROTEIN SELF-ASSOCIATION

5.2. For the dimerization reaction $2A \rightleftharpoons A_2$, the equilibrium constant in the mol L^{-1} scale is 10^5. What is the concentration in mol L^{-1} of dimer in equilibrium with 10^{-3} mol L^{-1} monomer?

SOLUTION

Rearranging Eq. (*5.1*), $[A_2] = K[A]^2$.
Thus, at equilibrium,

$$[A_2] = 10^5(10^{-3})^2 \text{ mol } L^{-1} = 0.1 \text{ mol } L^{-1}$$

5.3. For the dimerization reaction $2A \quad A_2$, in which A is a protein of molar weight 40,000 g mol^{-1}, the equilibrium constant in the mol L^{-1} scale is 10^6. Calculate the percentage by weight of dimer when the total concentration of the protein is 1 g L^{-1}.

SOLUTION

It is important to convert all units to a consistent set. Here, it is most convenient to use the mol L^{-1} scale. Thus, the molar concentrations are $[A] = c_A/M$ and $[A_2] = c_{A_2}/2M$, where c_A and c_{A_2} are the concentrations in g L^{-1} of A and A_2 respectively, and M is the molar weight of A.

Now, the total concentration of A is

$$c_T = c_A + c_{A_2}$$

Hence,

$$c_{A_2} = c_T - c_A$$

By substituting in Equation (5.1), we get

$$K = \frac{[A_2]}{[A]^2} = \frac{(c_T - c_A)/2M}{(c_A/M)^2}$$

Rearranging and then solving the quadratic in c_A gives two roots, one negative, and thus physically meaningless, and the other positive. The positive root is given by

$$c_A = \frac{-1 + (1 + 8Kc_T/M)^{1/2}}{4K/M}$$

Substituting for K, c_T, and M, we get

$$c_A = 0.13 \text{ g } L^{-1}$$

Therefore, the percentage by weight of monomer is 13, and that of the dimer is 87.

5.4. Repeat the above calculation for a total concentration of $10\,g\,L^{-1}$.

SOLUTION

Using the same procedure as in Prob. 5.3 and substituting into the expression for c_A, we get

$$c_A = \frac{-1 + [1 + (8 \times 10^6 \times 10)/(4 \times 10^4)]^{1/2}}{4 \times 10^6/4 \times 10^4} g\,L^{-1}$$

$$= \frac{-1 + (2{,}001)^{1/2}}{100} g\,L^{-1} = 0.437\,g\,L^{-1}$$

This represents only 4.4 percent of the total at this new concentration, and thus the percentage by weight of the dimer is 95.6 percent.

Note the general principle that can be drawn from this: As the total concentration of a self-associating molecule increases, the proportion of the associated form increases. A corollary is that dilution favors the monomeric species. This is an example of Ostwald's dilution principle.

HEMOGLOBIN

5.5. Fetal hemoglobin (HbF) is a tetramer of two α chains and two γ chains. The γ chains are similar to the β chains of HbA, but there are many sequence differences between them. One significant difference is that the residue H21 is the positively charged histidine in the β chain but the neutral serine in the γ chain. (*a*) Explain why HbF has a higher oxygen affinity than normal adult hemoglobin (HbA). (*b*) Why is this effect physiologically important?

SOLUTION

(*a*) The amino acid substitution at H21 affects the β-β cleft, where BPG is bound, there being two fewer positive charges in HbF. Thus, BPG is bound less strongly to HbF, and oxygen is more tightly bound and therefore is less readily released from HbF.

(*b*) This effect allows HbF to draw O_2 across the placenta from the maternal HbA for the use of the fetus.

5.6. In vertebrate hemoglobins, it is the tetramer that shows cooperativity on binding with oxygen; the dimer has no cooperative effect. The blood clam has been shown to contain two hemoglobins: a tetramer consisting of two pairs of unlike chains and a homodimer of a third, each of these chains having a tertiary structure similar to the globin fold. In the quaternary structure of the blood clam hemoglobins, the E and F helices, which lie on the proximal and distal sides of the heme group, are involved in subunit interactions forming an extensive contact between the chains in the dimer. This is "back to front" when compared with the vertebrate hemoglobins, where the E and F helices are on the outside of the molecule. Interestingly, both the dimer and the tetramer clam hemoglobins bind oxygen cooperatively. Explain how the mollusk dimer hemoglobin could demonstrate a cooperative effect.

SOLUTION

On combining with oxygen, the F helix, containing the proximal histidine residue, moves about 0.5 Å. In vertebrate hemoglobin, this movement, being on the surface of the molecule, has no direct effect, acting only on the tetrameric structure (the $\alpha_1\beta_1$ dimers are quite rigid). In the clam, such a movement will directly affect the structure of the dimer and may lead to an increase in the oxygen affinity, producing cooperativity.

5.7. There are many abnormal or mutant hemoglobins, some of which cause pathological conditions. One is sickle-cell hemoglobin (HbS), in which the glutamate residue in the sixth position of the normal human hemoglobin (HbA) β chain has been replaced by valine. This position, referred to as $\beta 6$, is on the outside of the hemoglobin molecule. Individuals who

are homozygous for HbS suffer from circulatory problems and anemia. This is because the red blood cells become sickle-shaped in the venous circulation, causing blockages in the capillaries, and are principally removed by the spleen. Why should such a small change, two amino acids in a total of 574, produce such a drastic effect?

SOLUTION

Amino acid residues on the outside of soluble proteins are almost without exception polar. The replacement of the strongly polar glutamate in HbA by the nonpolar valine in HbS leads to a hydrophobic area on the exterior of the molecule. Such a "sticky patch" would readily interact with a complementary binding site on another molecule. HbS, but not oxyHbS, has such a site, and this allows HbS to polymerize into long fiberlike chains that distort the red blood cell and produce the sickle shape. It is of interest that the HbS gene is common in tropical Africa, because the heterozygous individual, who rarely has sickling episodes, is protected against the most dangerous form of malaria. Should infection with malaria occur, the red cells would tend to sickle because of the increased oxidative stress imposed on them; these cells would then be removed from the circulation by the spleen.

5.8. Mutations in the α_1,β_2 sliding contact of hemoglobin generally cause impairment of the cooperative effect. In $Hb_{KEMPSEY}$, the aspartate residue G1 in the β chain is replaced by asparagine. This removes a hydrogen bond from the sliding contact (Fig. 5-12) that stabilizes the tense structure of deoxyhemoglobin but is not involved in the relaxed structure of oxyhemoglobin. What effect will this amino acid substitution have on the oxygen affinity of $Hb_{KEMPSEY}$?

SOLUTION

As the stability of the tense structure is reduced by the loss of this interaction in the two sliding contacts α_1,β_2, and α_2,β_1, less energy is required for oxygen to bind to $Hb_{KEMPSEY}$. Thus, $Hb_{KEMPSEY}$ has a higher oxygen affinity ($p50$ = 15 torr) and a lower Hill coefficient (see Chap. 9 and Prob. 5.21). Individuals with this hemoglobin compensate by producing more red blood cells.

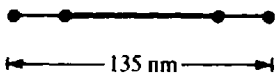
THE EXTRACELLULAR MATRIX

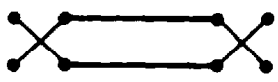
5.9. There is a continual turnover of collagen molecules in the body, e.g., in the remolding of bone and in wound healing. For many years it was believed that the triple helix was resistant to attack by all proteolytic enzymes. In the last 20 years, many mammalian collagenases have been found that cleave each of the three chains of the triple helix between a glycine and a large hydrophobic residue (e.g., leucine, isoleucine, or phenylalanine) roughly three-quarters of the distance from the N terminus. This is an unstable region, low in proline and hydroxyproline. Explain how this cleavage of the triple helix allows the degradation of collagen to proceed.

SOLUTION

The specific cleavage of collagen by mammalian collagenases produces two fragments that are unstable at body temperature and begin to uncoil into individual chains. These are then susceptible to attack by other proteolytic enzymes.

5.10. The monomers of type VI collagen are the shortest collagen molecules (150 nm), with only one-third of the mass being the triple-helical domain, the other two-thirds being the two terminal globular regions. On dispersion and rotary shadowing for the electron microscope, the following structures are the most abundant of those observed:





Explain the construction of the structures, and suggest how further assembly could occur to form extended microfibrils.

SOLUTION

The structure on the left shows a dimer with two antiparallel collagen type VI molecules assembled in a staggered manner, with the triple-helical segments overlapping by 75 nm. The dimers align symmetrically with the free (nonoverlapping) triple-helical ends crossing over like scissors to form the tetramer shown on the right. This tetramer then acts as the protomeric form, which aggregates end to end, with further crossing over of the ends to form the linear polymers of microfibrils.

5.11. The genes for the chains of fibrillar collagens (types I, II, III, and probably V) contain about 50 short exons (expressed regions), separated by long introns (nonexpressed regions); the introns are removed before the genetic information is translated into the protein chain (see Chap. 17). The exons for the triple-helical domain are unusual in that they all contain multiples of nine base pairs, 54 being the most common, followed by 108. If three nucleotide bases code for one amino acid, what is the significance of this gene organization?

SOLUTION

On translation each exon produces a length of oligopeptide that is a multiple of three amino acids; lengths of 18 and 36 amino acids are the most common, being equivalent to six or 12 turns of the triple helix. Each exon begins with a codon for glycine, and this is followed every nine nucleotides by another. It has been suggested that the collagen gene evolved from a small primordial gene and that the organization of the gene is now conserved to an extraordinary extent.

5.12. Defects in collagen synthesis can cause severe connective tissue disorders. In Ehlers-Danlos syndrome VII, those affected suffer from thin, velvety skin, which heals poorly, and from hypermobile joints and weak ligaments, a condition which leads to multiple joint dislocations. It has been reported that in this syndrome, exon 46 (numbering from the C-terminal end) is missing from the gene for the $\alpha 2(I)$ chain. This sequence includes the cleavage site for the procollagen N-peptidase and a lysine residue near the normal N terminus of the collagen $\alpha 2(I)$ chain. Describe the effect of the loss of this exon on the structure of type I collagen.

SOLUTION

The loss of the N-peptidase cleavage site would prevent the proper processing of the $\alpha 2(I)$ procollagen chain, and thus the mature chain would have a large N-terminal extension of about 50 amino acids, which would disrupt the proper assembly of the triple-helical collagen molecules into a fibril. The lack of the lysine residue would prevent the formation of one of the four major interchain cross-links that the $\alpha 2(I)$ chain makes with its quarter-staggered neighbors. The overall effect is to weaken tissues that depend on type I collagen for tensile strength.

5.13. How many specific glycosyltransferases are required for the sequential addition of monosaccharide units in the synthesis of chondroitin sulfate? Note that substrate specificity is strict.

SOLUTION

Six glycosyltransferases are required (Fig. 5-21) and they act in the following order: (1) xylyltransferase, (2) galactyltransferase I (specific for xylose), (3) galactyltransferase II (specific for galactose), (4) glucuronate transferase I (specific for Gal-Gal), (5) *N*-acetylglucosamine transferase, (6) glucuronate transferase II (specific for *N*-acetylglucosamine). Finally, a sulfotransferase will catalyze the addition of the sulfate groups.

CYTOSKELETON

5.14. If the spectrin heterodimer can associate only head to head, explain the observation of higher polymers as large as the dodecamer.

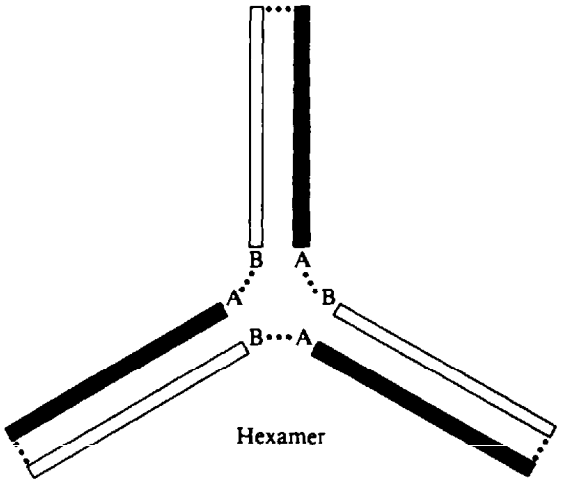


Fig. 5-39

SOLUTION

The head of each spectrin chain interacts with the head of the complementary chain of another heterodimer. In the tetramer, there are paired interactions (Fig. 5-31), while in higher oligomers, a closed loop is formed, as the head region is quite flexible. This is shown for the hexamer in Fig. 5-39, but this self-association may continue indefinitely.

5.15. Several cytoskeletal proteins contain long coiled-coil domains in which two or even three α helices coil around each other. What sequence criteria would be necessary for the α helices to form stable coiled coils in aqueous solutions?

SOLUTION

The α helix contains 3.6 amino acid residues per turn, and in a coiled coil, this produces seven residues for every two turns of the α helix. If these seven residues are denoted a, b, c, d, e, f, and g, then residues a and d (with e as an alternative) would be expected to be hydrophobic. These residues would be next to each other on one side of the α helix and in the center of the coiled coil, which is thus stabilized by hydrophobic interactions. This is known as the *heptad repeat*.

5.16. Many assemblies of actin and myosin in cells are temporary structures. One example is the beltlike *contractile ring* that appears during cell division just below the plasma membrane. As the ring contracts, the center of the cell is constricted until two daughter cells are produced. What interactions and processes must occur for this constriction to take place, noting that the ring remains a constant thickness?

SOLUTION

To allow the force generated between the actin and myosin to act upon the membrane, the contractile ring must be anchored to the membrane by specific actin-binding proteins. As the cell constricts, the contractile ring must disassemble little by little to remain a constant thickness.

5.17. Explain how the alkaloid drug *colchicine*, which binds tightly to tubulin dimers, blocks mitosis.

SOLUTION

The binding of the tubulin dimer by colchicine not only will prevent polymerization and the formation of microtubules but also, by reducing the concentration of free tubulin below the critical concentration,

will cause depolymerization and breakdown of the mitotic spindle. Similar drugs have been used as antitumor agents.

5.18. The nuclear envelope is a supramolecular assembly that lines the inner surface of the nucleus. The major proteins of the envelope are the A, B, and C *lamins*, and it has been suggested that they may be members of the intermediate filament family. What characteristic features should the lamins possess for this to be so?

SOLUTION

The intermediate filament family consists of proteins with a large central region with considerable sequence homology, which folds up into an α helix, with the heptad repeat allowing the formation of coiled-coil domains. The positions of the short random sections between the coiled-coil domains are conserved. Strong sequence divergence in the two globular terminal regions would be expected.

Supplementary Problems

5.19. For the dimerization reaction $2A \rightleftharpoons A_2$, calculate the concentration in mol L^{-1} of dimer for the following conditions:

	K	[A] (mol L^{-1})
(*a*)	10^3	10^{-3}
(*b*)	10^4	10^{-2}
(*c*)	10^6	10^{-4}

5.20. For the dimerization reaction $2A \rightleftharpoons A_2$, in which A is a protein of molar weight 50,000 g mol^{-1}, calculate the percentage by weight of the dimer when the total concentration of the protein is 1 g L^{-1} and the equilibrium constant in the mol L^{-1} scale is (*a*) 10,000; (*b*) 10^8.

5.21. Early this century, A.V. Hill derived a useful equation that describes the oxygen dissociation curve of hemoglobin fairly accurately. It is

$$Y = \frac{(pO_2)^n}{(pO_2)^n + (p50)^n}$$

where Y is the fraction of hemoglobin oxygen-binding sites occupied by oxygen, pO_2 the partial pressure of O_2, $p50$ the partial pressure of oxygen when $Y = 0.5$ (i.e., 50 percent of the sites are filled), and n the *Hill coefficient*, which is a measure of cooperativity. The value of n for human HbA is 2.8, while that of myoglobin, which is not cooperative, is 1. Calculate the change in saturation ΔY for the situation in which hemoglobin goes from lung to working muscle and in which pO_2 for lung is 100 torr and pO_2 for muscle is 20 torr under the following circumstances:

	Protein	$p50$ (torr)	n
(*a*)	HbA	26	2.8
(*b*)	Noncooperative Hb	26	1
(*c*)	Mb-like Hb	1	1
(*d*)	HbA with high DPG	31	2.8
(*e*)	Clamlike dimer Hb	26	1.5

5.22. One of the salt links stabilizing deoxyhemoglobin involves a chloride ion. Explain how this ion could act as an allosteric effector.

5.23. In the abnormal hemoglobin Hb_{KANSAS}, residue G4 on the β chain is threonine instead of the asparagine in HbA. This change weakens the interactions in the sliding contact in the relaxed state of oxyhemoglobin without affecting the tense state. How would the oxygen affinity of Hb_{KANSAS} compare with that of HbA?

5.24. Which of the following synthetic polypeptides would be expected to form a triple helix similar to collagen?

(*a*) $(\text{Pro-Gly})_n$

(*b*) $(\text{Pro-Pro-Gly})_n$

(*c*) $(\text{Pro-Gly-Gly})_n$

(*d*) $(\text{Pro-Gly-Pro-Pro})_n$

Of those which form a triple helix, which would be the most stable?

5.25. Why is the triple helix of collagen a better structure for providing tensile strength than the α helix?

5.26. The following is a sequence of part of a pro-α-chain of collagen: -Gly-Pro-Met-Gly-Pro-Ser-Gly-Pro-Arg-Gly-Leu-Pro-Gly-Pro-Pro-Gly-Ala-Pro-Gly-Pro-Gln-Gly-Pro-Arg-Gly-Pro-Pro-Gly-Glu-Pro-Gly-Glu-Pro-Gly-Ala-Ser-. (*a*) How many residues would be hydroxylated by the enzyme *prolylhydroxylase*? (*b*) How many peptides would be produced after digestion by the collagenase from one of the gas gangrene organisms *Clostridium histolyticum*? The specificity of the enzyme is to cleave the bond before glycine in the sequence -Pro-X-Gly-Pro-.

5.27. In the calcification process of bone formation, the initial crystals of hydroxyapatite are found at intervals of 67 nm along the collagen fiber. What is the reason for this?

5.28. In osteoarthritis, fragments of proteoglycans are produced. It is believed that those fragments are formed by the action of proteases produced in response to inflammation. Where is the action of the proteases on the proteoglycan most likely to occur?

5.29. In a 0.01% solution, the domains of hyaluronate occupy the total volume. What would be the ratio of the volume occupied by one molecule of hyaluronate ($M_r = 1.5 \times 10^6$) to the volume occupied by five molecules of collagen ($M_r = 300{,}000$; diameter 1.5 nm; length 300 nm)?

5.30. How might the identification of the type and subtype of intermediate filament help in the diagnosis and classification of cancer?

5.31. In fast axonal transport, dynein motors serve to bring vesicles from near the end of the axon (the plus end of the microtubules) toward the cell body (the minus end of the microtubules). The distance traveled can be as much as a meter. Since dynein is a minus-end directed motor, how does it get out to the plus ends of the microtubules in the first place?

Chapter 6

Lipids, Membranes, Transport, and Signaling

6.1 INTRODUCTION

Lipids are defined as water-insoluble compounds extracted from living organisms by weakly polar or nonpolar solvents. This definition is based on a *physical* property, in contrast to the definitions of proteins, carbohydrates, and nucleic acids, which are based on chemical structure. Consequently, the term *lipid* covers a structurally diverse group of compounds, and there is no universally accepted scheme for classifying lipids.

EXAMPLE 6.1

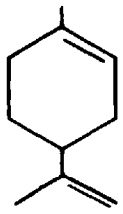
Consider the following compounds, most of which are lipids:

$CH_3(CH_2)_{14}COO^-$
(1) **Palmitate**

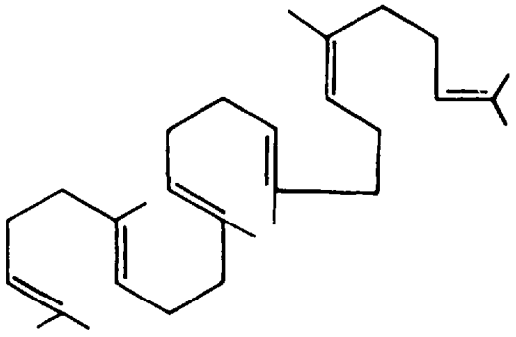
$CH_3CH_2COO^-$
(2) **Propionate**

$CH_3(CH_2)_{14}CH_2OH$
(3) **Cetyl alcohol**

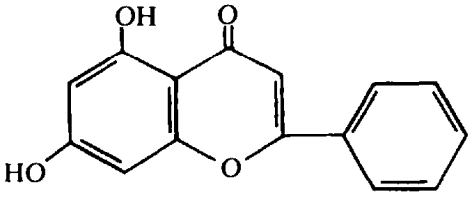
(4) **Benzene**



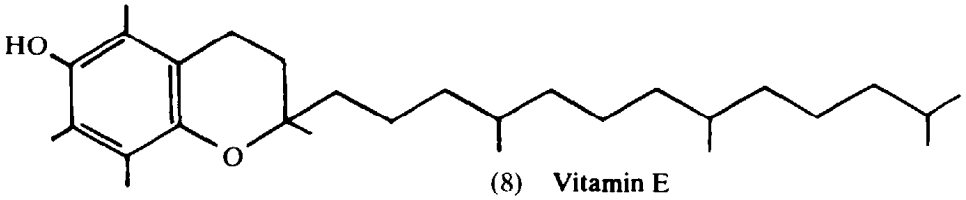
(5) **Limonene**



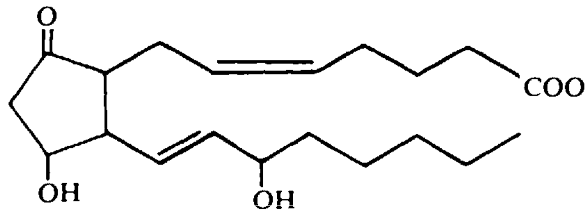
(6) **Squalene**



(7) Chrysin



(8) **Vitamin E**



(9) Prostaglandin E_2

Compounds 1, 3, and 5 to 9 are lipids because they are biological in origin and are soluble in organic solvents. The latter property arises because they contain a *high proportion of carbon and hydrogen* and are therefore insoluble in water. Compound 4 is not a lipid because it does not occur free in living organisms. Compound 2 is water-soluble, but because it is a member of the same series of compounds as compound 1, it is usually considered to be a lipid.

6.2 CLASSES OF LIPIDS

A common feature of all lipids is that, biologically, their hydrocarbon content is derived from the polymerization of acetate followed by *reduction* of the chain so formed. (However, this process also occurs for the synthesis of some compounds that are not lipids and therefore cannot be used as a definition of lipids.) For example, polymerization of acetate can give rise to the following:

1. Long, linear hydrocarbon chains:

$$n\mathrm{CH_3COO^-} \rightarrow \rightarrow \mathrm{CH_3COCH_2CO} \cdots \rightarrow \rightarrow \mathrm{CH_3CH_2CH_2CH_2} \cdots$$

   The products are *fatty acids*, $CH_3(CH_2)_nCOOH$, which in turn can give rise to amines and alcohols. Lipids containing fatty acids include the *glycerolipids*, the *sphingolipids*, and *waxes*.

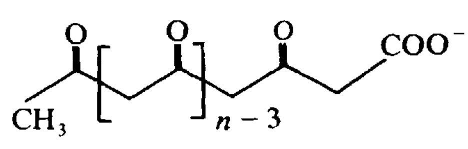
2. Branched-chain hydrocarbons via a five-carbon intermediate, isopentene (*isoprene*):

$$3\mathrm{CH_3COO} \longrightarrow \longrightarrow (\mathrm{CH_3{-}CH{=}\overset{\displaystyle CH_3}{\overset{|}{C}}{-}CH_2{-}}) \longrightarrow \longrightarrow \textit{terpenes}$$

   (including *steroids* and *carotenoids*)

3. Linear or cyclic structures that are only partially reduced:

$$n\mathrm{CH_3COO^-} \longrightarrow \longrightarrow$$



   These are called *acetogenins* (or sometimes *polyketides*). Many of these compounds are aromatic, and their pathway of formation is the principal means of synthesis of the benzene ring in nature. Not all are lipids, because partial reduction often leaves oxygen-containing groups, which render the product soluble in water.

EXAMPLE 6.2

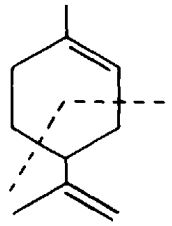
For the lipids in Example 6.1, the routes of synthesis (just discussed) are:

Compounds 1, 3, and 9 are made via route (1).

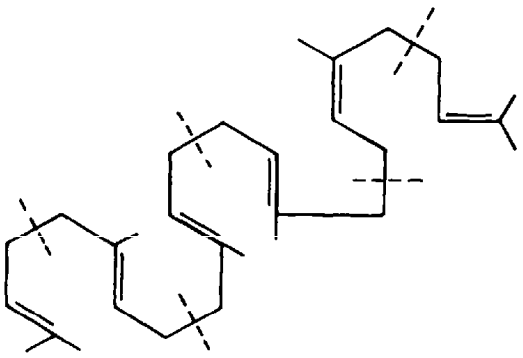
Compounds 5 and 6 are made via route (2).

Compound 7 is made by route (3), and compound 8 has a mixed origin via routes (2) and (3).

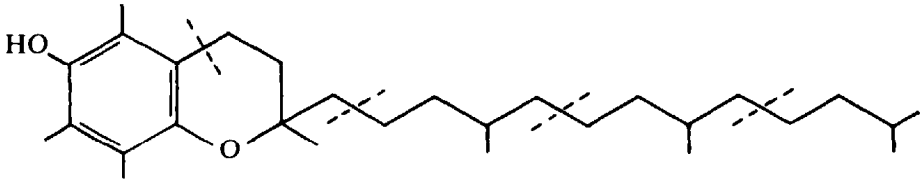
For those lipids synthesized by route (2), the ways in which the isoprene units are linked are shown by the broken lines in the following structural formulas:



Limonene



Squalene



Vitamin E

6.3 FATTY ACIDS

Over 100 fatty acids are known to occur naturally. They vary in chain length and degree of unsaturation. Nearly all have an even number of carbon atoms. Most consist of linear chains of carbon atoms, but a few have branched chains. Fatty acids occur in very low quantities in the free state and are found mostly in an *esterified* state as components of other lipids. The pK_a of the carboxylic acid group is about 5, and under physiological conditions, this group will exist in an ionized state called an *acylate* ion; e.g., the ion of palmitic acid is *palmitate*, $CH_3(CH_2)_{14}COO^-$.

EXAMPLE 6.3

The following are some biologically important fatty acids.

(*a*) Saturated:

$CH_3(CH_2)_{14}(COOH)$	$CH_3(CH_2)_{16}COOH$
Palmitic acid	Stearic acid
(hexadecanoic acid)	(octadecanoic acid)
16:0	18:0

(*b*) In *unsaturated* fatty acids, the double bond nearly always has the cis conformation:

$CH_3(CH_2)_5CH{=}CH(CH_2)_7COOH$
Palmitoleic acid
(*cis*-9-hexadecenoic acid)
$16{:}1^{\Delta 9}$

(*c*) In *polyunsaturated* fatty acids, the double bonds are rarely conjugated:

$$CH_3(CH_2)_4CH{=}CH{-}CH_2{-}CH{=}CH(CH_2)_7COOH$$

Linoleic acid
(*cis*,*cis*-9,12-octadecadienoic acid)
$18{:}2^{\Delta 9,12}$

A *number notation* used widely for indicating the structure of a fatty acid is shown under the names of the fatty acids in Example 6.3. To the left of the colon is shown the number of C atoms in the acid; to the right, the number of double bonds. The position of the double bond is shown by a superscript Δ followed by the number of carbons between the double bond and the end of the chain, with the carbon of the carboxylic acid group being called 1.

EXAMPLE 6.4

The number notations, systematic name, and trivial name, respectively, for three more fatty acids are:

$$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$$

$18{:}1^{\Delta 9}$ *cis*-9-octadecenoic acid (oleic acid)

$$CH_3CH_2CH{=}CH{-}CH_2{-}CH{=}CH{-}CH_2{-}CH{=}CH(CH_2)_7COOH$$

$18{:}3^{\Delta 9,12,15}$ all *cis*-9,12,15-octadecatrienoic acid (α-linolenic acid)

$$CH_3(CH_2)_4CH{=}CH{-}CH_2{-}CH{=}CH{-}CH_2{-}CH{=}CH{-}CH_2{-}CH{=}CH(CH_2)_3COOH$$

$20{:}4^{\Delta 5,8,11,14}$ all *cis*-5,8,11,14-eicosatetraenoic acid (arachidonic acid)

The melting points of different fatty acids differ markedly. For example:

Palmitic acid, 63°C; stearic acid, 70°C; oleic acid, 13°C

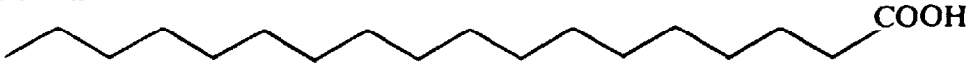
Elaidic acid (*trans*-9-octadecanoic acid), 44°C

Linoleic acid, −5°C; α-linolenic acid, −11°C

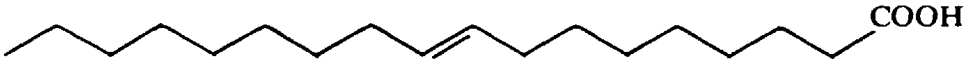
Question: Why do differences in melting point exist between fatty acids containing the same number of carbon atoms?

The preferred conformation of a chain of saturated C atoms is a long, straight structure. A cis double bond will cause a bend in the structure, making it less likely to pack into a crystal than will a saturated molecule of the same length. A trans double bond does not cause a bend in the chain. For example:

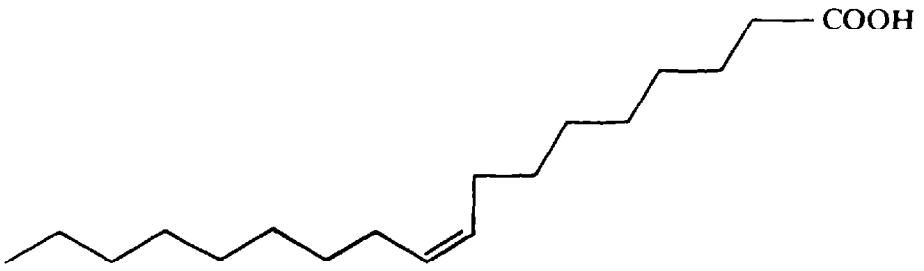
1. Saturated chain



2. Chain with one trans double bond



3. Chain with one cis double bond



Straight molecules can pack together more densely and give crystals of higher melting point than the melting points of bent molecules of the same size; in other words, more energy is required to separate the molecules when they are heated.

The presence of cis rather than trans double bonds in unsaturated fatty acids ensures that lipids containing fatty acids have low melting points and are therefore fluid at physiological temperatures.

Question: Apart from unsaturation, what other structural feature of a fatty acid could affect its melting point?

Branching. For example, stearic acid (18 carbons) and arachidic acid (20 carbons) are both saturated and linear. They melt at 70°C and 75°C, respectively, whereas 10-methylstearic acid melts at 10°C. However, branched fatty acids are rare in living systems and the evolution of synthetic pathways for them has not occurred as a major device for keeping lipids fluid.

$$CH_3(CH_2)_7\underset{\displaystyle CH_3}{\underset{|}{C}}H(CH_2)_8COOH$$

10-Methylstearic acid

6.4 GLYCEROLIPIDS

Glycerolipids are lipids containing *glycerol* in which the hydroxyl groups are substituted in some way. In terms of quantity, these are by far the most abundant lipids in animals. Somewhat similar in structure, but occurring at levels of less than 1 percent of the glycerolipids, are lipids containing *diols*, e.g., ethylene glycol (ethane diol) and 1,2- and 1,3-propanediol. Because of their rarity, lipids based on diols will not be discussed further.

$$\begin{array}{r} CH_2OH \\ | \\ HOCH \\ | \\ CH_2OH \end{array} \qquad \begin{array}{l} CH_2OH \\ | \\ CH_2OH \end{array} \qquad \begin{array}{r} CH_2OH \\ | \\ HOCH \\ | \\ CH_3 \end{array}$$

Glycerol — Ethylene glycol — 1,2-Propanediol

Triacylglycerols

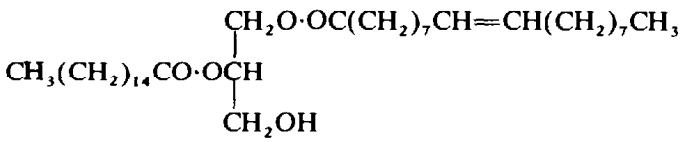
Triacylglycerols (TAGs) are *neutral* glycerolipids and are also known as *triglycerides*. In the TAGs the three hydroxyl groups of glycerol are each esterified, usually by *different* fatty acids. This makes the second carbon of the glycerol moiety chiral. A special convention is used for dealing with the naming of TAGs and other glycerol derivatives. The derivative is drawn as a Fischer projection (Chap. 2) with the *secondary hydroxyl to the left*, and the carbon atoms are numbered 1, 2, and 3 from the top. The prefix *sn-* (for *stereospecifically numbered*) precedes the name *glycerol*. For example,

$$\begin{array}{rl} & CH_2O{\cdot}OCR \\ & | \\ R'CO{\cdot}O— & C—H \\ & | \\ & CH_2O{\cdot}OCR'' \end{array}$$

1,2,3-Triacyl-*sn*-glycerol

EXAMPLE 6.5

The structure of 1-oleoyl-2-palmitoyl-*sn*-glycerol is



This is a *diacylglycerol*. Diacyl- and monoacylglycerols are found in cells, but only in small amounts; they are metabolites of TAGs and phospholipids.

TAGs are the most abundant lipids found in animals. This is because they function as a food store. Although found in most cells, TAGs are particularly present in the cells of adipose tissue where they form *depot fat*. The hydrolysis of the ester bonds of TAGs and the release of glycerol and fatty acids from adipose tissue is called *fat mobilization*. Depot fat is a water-free mixture of TAGs differing from each other in the nature of the three fatty acyl groups from which they are built.

Question: Depot fat has a relatively high content of unsaturated fatty acids. What advantage does this have for the cell?

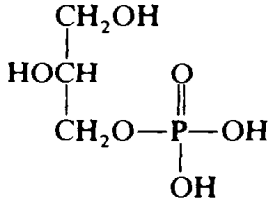
The fat exists in a liquid state. Solid fat would present only a small surface area for enzymes in the cytoplasmic water that mobilize it. Also, solid fat would render the adipose tissue rigid and unyielding during mechanical stress.

Usually the melting point of depot fat is only a few degrees below body temperature. The fatty acid composition of depot fat thus is a compromise between the requirement for keeping the fat liquid and the ability to store as much energy as possible. (Unsaturated fatty acids yield less energy than saturated ones of the same size, when oxidized.)

Phosphoglycerides

Phosphoglycerides are *polar* glycerolipids and are often referred to as *phospholipids*. However, some other lipids, not containing glycerol, also contain phosphorus, and the term phospholipid also describes these. The term *phospholipid* will be used here to describe *any* lipid containing phosphorus.

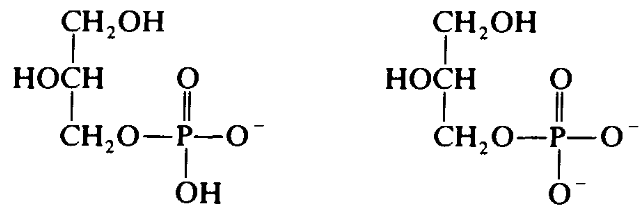
All phosphoglycerides are derived from *sn*-glycerol-3-phosphoric acid in which the phosphoric acid moiety is esterified with certain alcohols and the hydroxyl groups on C-1 and C-2 are esterified with fatty acids.



sn-Glycerol-3-phosphoric acid

Question: In what form will *sn*-glycerol-3-phosphoric acid exist at physiological pH?

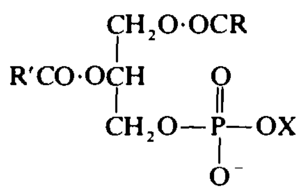
The pK_a values of phosphomonoesters are ~1 and ~6; therefore, there will be two ionic species at pH 7, with the dianion predominating:

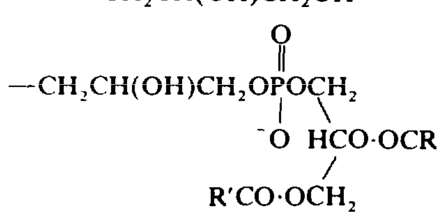
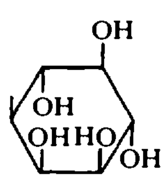


The phosphoglycerides are named and classified according to the nature of the alcohol esterifying the glycerol phosphate (Table 6.1).

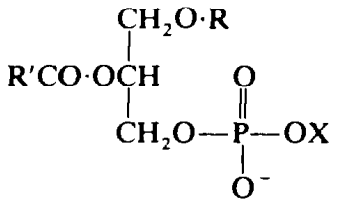
In most phosphoglycerides, the fatty acid substituted on C-1 is saturated and that on C-2 is unsaturated. Although phosphoglycerides are referred to in the singular (as in Table 6.1), they are mixtures in which compounds with the same *X* group are esterified with a variety of different fatty acids. In some instances C-1 is *etherified* (not esterified) with a long-chain fatty *alcohol*. This

Table 6.1. Some Major Classes of Phosphoglyceride



Name of *X*—OH	Structure of *X*	Name of Phosphoglyceride (Symbol)
Water	—H	Phosphatidate [ion of phosphatidic acid] (PA)
Ethanolamine	$—CH_2CH_2\overset{+}{N}H_3$	Phosphatidylethanolamine (PE)
Choline	$—CH_2CH_2\overset{+}{N}(CH_3)_3$	Phosphatidylcholine (PC)
Serine	$—CH_2CH(COO^-)\overset{+}{N}H_3$	Phosphatidylserine (PS)
Glycerol	$—CH_2CH(OH)CH_2OH$	Phosphatidylglycerol (PG)
Phosphatidylglycerol		Diphosphatidylglycerol [cardiolipin] (DPG)
Inositol		Phosphatidylinositol (PI)

phosphoglyceride is known as a *plasmalogen*:

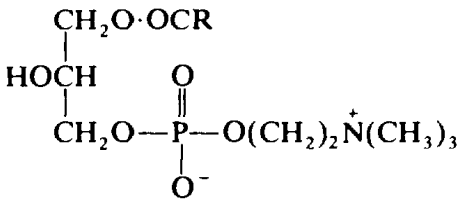


Question: Why are phosphoglycerides often described as *polar* lipids?

The term arises because of the charges on the moiety bearing the esterified phosphate. The term *polar* is used in a relative sense; i.e., relative to TAGs, the phosphoglycerides are polar. But in an *absolute sense*, they are nonpolar overall and insoluble in water.

About 1 percent of the total phosphoglycerides that occur in animal cells are in the form of *lysophosphoglycerides*, in which one of the acyl substituents, usually from C-2, is missing. The lysophosphoglycerides are named by adding the prefix *lyso-* to the name of the parent phosphoglyceride.

Question: What is the structural formula for *lysophosphatidylcholine*?

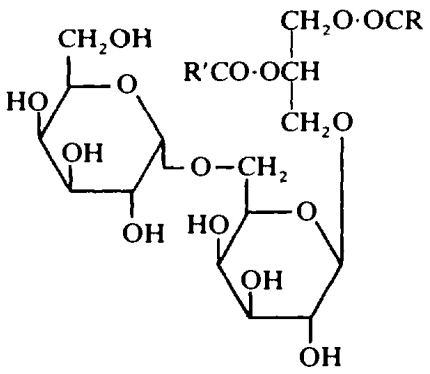


Glycoglycerolipids

Glycoglycerolipids are similar in some respects to the phosphoglycerides; namely, they have hydrophobic and polar (hydrophilic) parts, the latter being provided by a carbohydrate moiety rather than by an esterified phosphate.

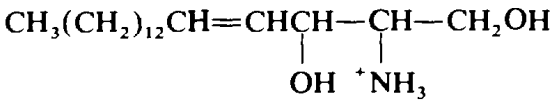
EXAMPLE 6.6

A typical glycoglycerolipid is 6-α-D-galactopyranosyl-β-D-galactopyranosyldiglyceride. The method for writing the carbohydrate portion is explained in Chap. 2, and the structure of the lipid is:

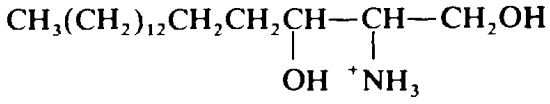


6.5 SPHINGOLIPIDS

Sphingolipids are built from long-chain, hydroxylated bases rather than from glycerol. Two such bases are found in animals: *sphingosine* and dihydrosphingosine (*sphinganine*), with the former being much more common.



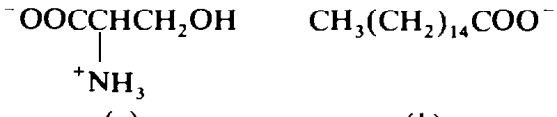
Sphingosine



Sphinganine

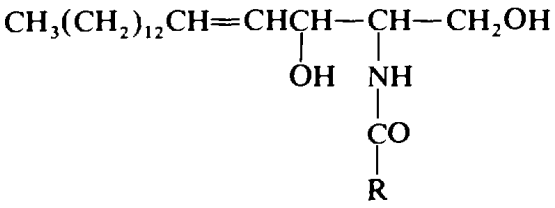
Question: From which compounds is sphinganine synthesized in living systems?

Notice the polar part is related to the amino acid *serine* (*a*) and the nonpolar part resembles *palmitate* (*b*). A reaction between these two compounds, with the elimination of CO_2, is followed by a reduction reaction which yields sphinganine.



(*a*) (*b*)

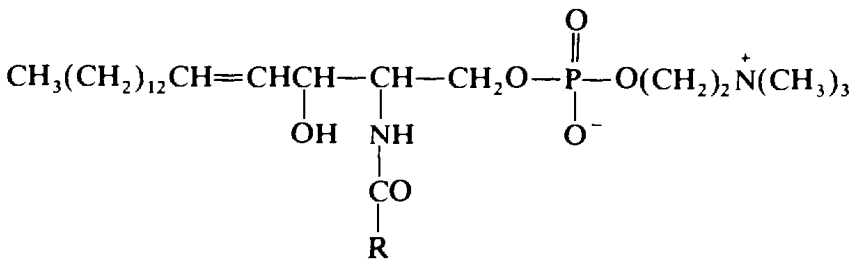
When the amino group of sphingosine or sphinganine is acylated with a fatty acid, the product is a *ceramide*,



The primary hydroxyl group is substituted in one of two ways to give two classes of sphingolipid; these are the phosphosphingolipids and glycosphingolipids.

Phosphosphingolipids

In *phosphosphingolipids*, the primary hydroxyl group is esterified with choline phosphate. The lipid is known as *sphingomyelin*. It has the structure:



Sphingomyelin

The fatty acyl (R) groups found in sphingomyelin are unusual; e.g.:

Lignoceric 24:0

Nervonic $24{:}1^{\Delta 15}$

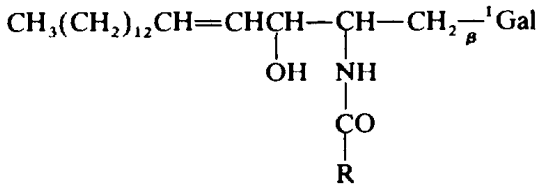
Cerebronic 24:0, C-2 hydroxylated

Glycosphingolipids

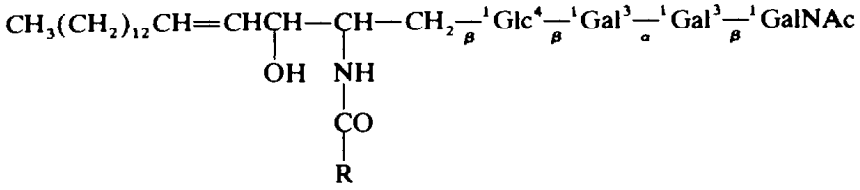
In *glycosphingolipids*, the primary hydroxyl group is glycosylated, i.e., substituted with a carbohydrate, either a mono- or an oligosaccharide. Glycosphingolipids that contain the sugar *sialic acid* in the carbohydrate portion are called *gangliosides*. At least 50 types of glycosphingolipid are known, based on differences in the carbohydrate portion of the molecule. In addition, each type displays variation in the types of fatty acid found in the ceramide portion.

EXAMPLE 6.7

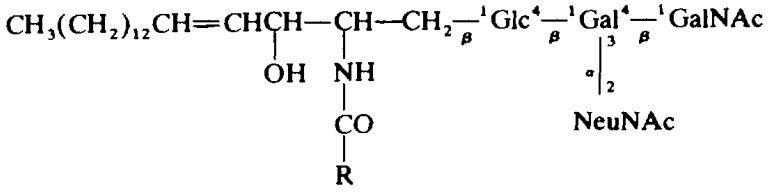
Some common glycosphingolipids are:



Galactosylceramide



Globoside



Tay-Sachs ganglioside

Note: Glc = glucosyl; Gal = galactosyl; GalNAc = *N*-acetylgalactosaminyl; NeuNAc = sialyl.

6.6 LIPIDS DERIVED FROM ISOPRENE (TERPENES)

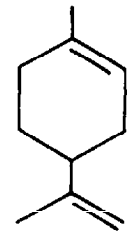
The name *terpene* was applied originally to the steam-distillable oils obtained from turpentine (an extract of pine). It was recognized that:

1. Most of the compounds present in the oil have the formula $C_{10}H_{15}$.
2. Terpenes with more than 10 carbons exist, the number of carbons being usually a multiple of five. The structures are extraordinarily diverse.

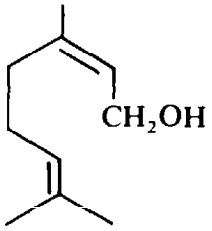
3. Many similar water-insoluble compounds are distributed very widely; particularly large quantities are found in many plants, but they exist also in most other living organisms.

Terpenes

Terpenes with 10 C atoms are known as *monoterpenes*. Two examples are:

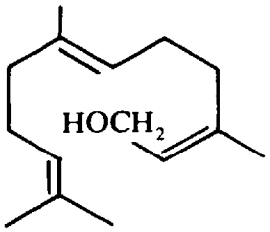


Limonene

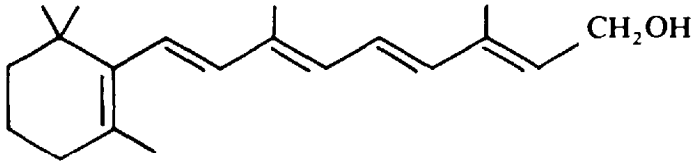


Geraniol

Sesquiterpenes and *diterpenes* have 15 and 20 C atoms, respectively. Two examples are:

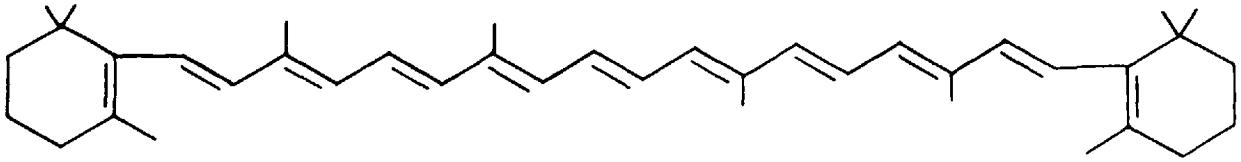


Farnesol



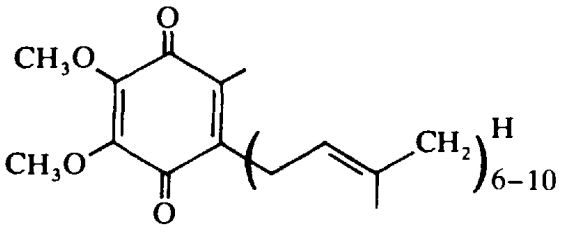
Vitamin A

Triterpenes (30 C atoms) and *tetraterpenes* (40 C atoms) require special attention because they give rise to *steroids* and *carotenoids*. Squalene (a triterpene) is compound 6 in Example 6.1.

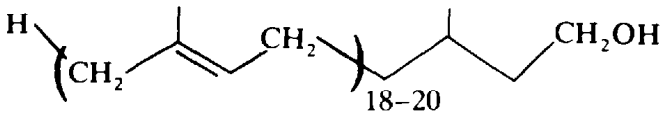


β-Carotene

Polyisoprenoid compounds exist, e.g., rubber, but in a biochemical context, the *ubiquinones* and *dolichols* are particularly important (see Chap. 14).



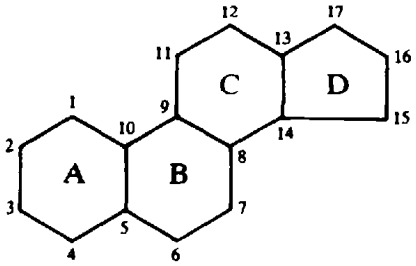
Ubiquinone



Dolichol

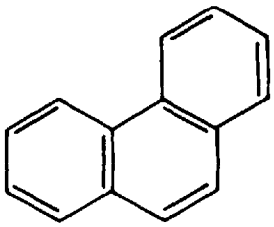
Steroids

Structurally, steroids are derivatives of the reduced aromatic hydrocarbon *perhydrocyclopentanophenanthrene*.



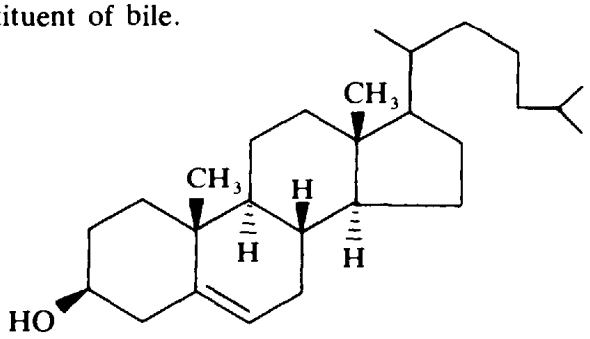
Perhydrocyclopentanophenanthrene

Although related structurally to phenanthrene, these compounds are *not acetogenins* but are true terpenes that are synthesized in living systems from isoprene via squalene. (Compare the structure of squalene—compound 6 in Example 6.1—with that of perhydrocyclopentanophenanthrene.)

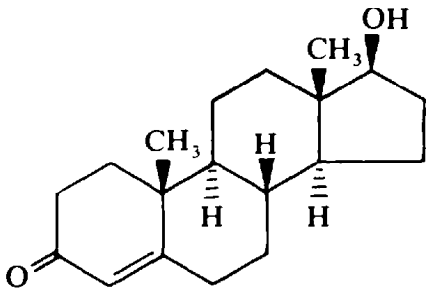


Phenanthrene

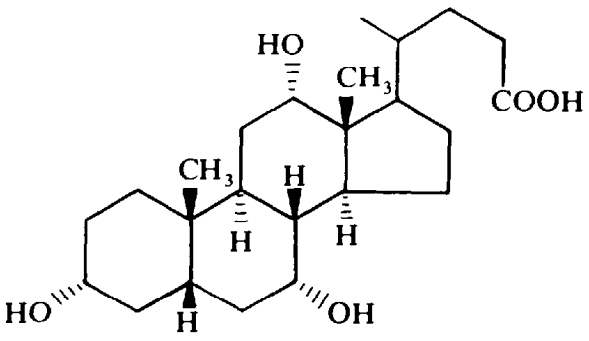
Sterols are steroids containing one or more hydroxyl groups. Some examples are *cholesterol*, a component of the cytoplasmic membrane of animal cells, *testosterone*, a hormone, and *cholic acid*, a constituent of bile.



Cholesterol

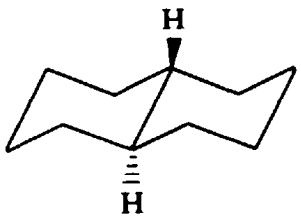


Testosterone

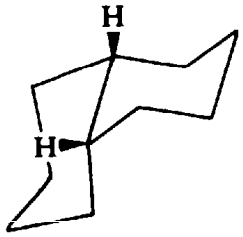


Cholic acid

The fused-ring system is an essentially planar structure. Bonds shown ► and ||||, indicate the substituents that are, respectively, in *front* (β) and *behind* (α) the general plane of the sterol. The configuration of substituents of the two carbon atoms shared between two fused rings determines whether the rings are fused in the cis or trans conformation. Trans-fused rings are flat structures, but cis-fused rings are bent structures. Cholesterol and testosterone have trans-fused rings and are essentially planar structures.

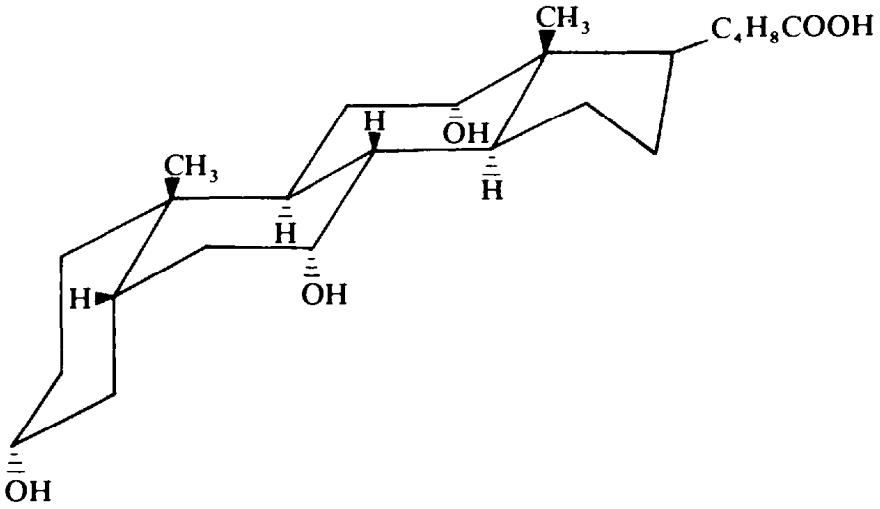


Trans-fused rings



Cis-fused rings

Question: What is the three-dimensional arrangement of the rings of cholic acid?



Carotenoids

These are hydroxylated derivatives of the 40-carbon hydrocarbons called *carotenes*. Because these compounds are highly conjugated, they absorb visible light; most of the yellow and red pigments occurring naturally are carotenes and carotenoids. These pigments are often involved with the interaction of living systems with light. Thus in animals, *β-carotene* (a tetraterpene) is metabolically converted to *vitamin A* (both structures are shown earlier in this section), which in turn is necessary for visual activity.

6.7 BEHAVIOR OF LIPIDS IN WATER

By definition, lipids are insoluble in water. Yet they exist in an aqueous environment, and their behavior toward water is therefore of critical importance biologically.

Many types of lipid are said to be *amphiphilic*, meaning they consist of two parts—a nonpolar hydrocarbon region and a region that is polar, ionic, or both. (The term *amphiphilic* has tended to replace *amphipathic*, used formerly.)

Question: Which of the following lipids are amphiphilic: fatty acids; acylate ions; TAGS; cholesterol; phosphoglycerides; phosphosphingolipids; glycosphingolipids?

Fatty acids, TAGs, and cholesterol are not amphiphilic; what polarity they have is extremely weak. All the others possess at least one *formal charge* or an abundance of hydroxyl groups in one part of the molecule.

When amphiphilic molecules are dispersed in water, their hydrophobic parts (i.e., hydrocarbon chains) segregate from the solvent by self-aggregation. The aggregated products are known as *micelles* (for those aggregates dispersed in water) and *monolayers* (for those aggregates at the water-air boundary). A diagram may be drawn (see Fig. 6-1) showing how an amphiphile (symbolized ○—, where ○ represents the polar head and — represents the hydrocarbon tail) will form a monolayer on the surface of water. The polar heads are in contact with the polar water, thus ensuring that the nonpolar tails are as remote as possible from water.

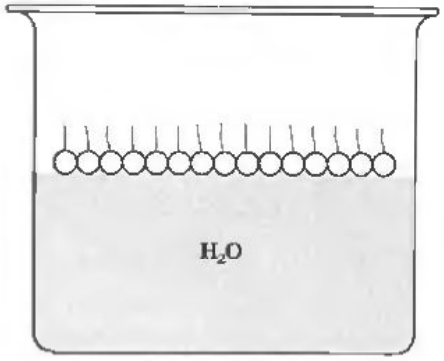


Fig. 6-1

The tendency for hydrocarbon chains to become remote from the polar solvent, water, is known as the *hydrophobic effect* (Chap. 4). Hydrocarbons form no hydrogen bonds with water, and a hydrocarbon surrounded by water facilitates the formation of hydrogen bonds between the water molecules themselves. The bulk water is more structured than it is in the absence of the hydrocarbon; i.e., it has lost entropy (Chap. 10) and is thus in a thermodynamically less favorable state. This state is obviated by the hydrocarbon being organized so that it is remote from water, thus rendering the water molecules near to it less ordered. Thus the hydrophobic effect is said to be *entropically driven*.

Only a small quantity of an amphiphilic lipid dispersed in water can form a monolayer (unless the water is spread as a very thin film), in which case the bulk of the lipid will form soluble micelles. Micelles can take a variety of forms, each satisfying the hydrophobic effect. Fig. 6-2 shows one such form, representing a spherical micelle, although ellipsoidal, diskoidal, and cylindrical variations are possible.

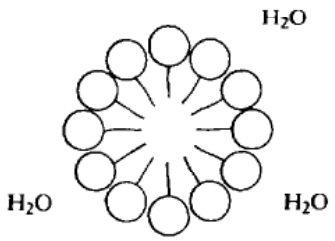
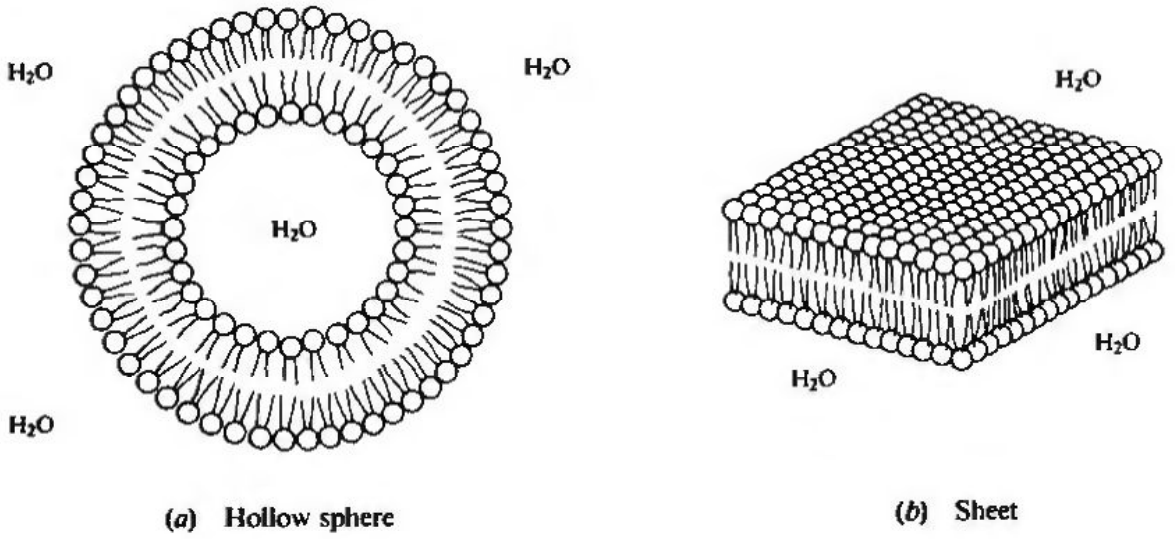


Fig. 6-2



(*a*) Hollow sphere (*b*) Sheet

Fig. 6-3 Forms of lipid bilayers.

Question: Are there other forms that amphiphilic lipids can adopt in water?

A bilayered structure in the form of a closed, hollow sphere is also possible [Fig. 6-3(*a*)]. This type of structure is called a *vesicle*. The primary concept of a vesicle is two sheets of lipid with their hydrocarbon chains opposed [a *bilayer*, Fig. 6-3(*b*)]. An isolated bilayer cannot exist as such in water, because exposed hydrocarbon tails would exist at the edges of the sheet; however, this situation is obviated by the sheet's curving to form a self-sealed, hollow sphere.

Both micelles and bilayers arise through the operation of two opposing forces: (1) attractive forces between hydrocarbon chains (van der Waals forces) caused by the hydrophobic effect forcing such chains together and (2) repulsive forces between the polar head groups.

Question: What determines the lower limit of micellar size?

The hydrophobic effect does; a minimum number of hydrocarbon chains must associate before the water-hydrocarbon interface is eliminated. This association process is a cooperative one, and the micelles therefore have a minimum size.

Question: What determines the upper limit of micellar size?

The repulsion of the polar heads does. If there are *two* hydrocarbon chains per polar head group, the nonpolar volume per head group is twice that of an amphiphilic lipid with one hydrocarbon chain. The greater repulsive force in the latter prevents the lipid molecules from coming too close and thus keeps the micellar size small. The weaker repulsive force and larger hydrocarbon volume in the former allow very much larger structures to form; namely, bilayers and vesicles.

Question: Which lipids form small micelles and which form vesicles and bilayers?

The length of the hydrocarbon chain relative to the size of the polar head group of an amphiphile influences whether it forms micelles or vesicles in polar solvents such as water. The answers are acylate ions and lysophosphoglycerides form small micelles; phosphoglycerides, phosphosphingolipids, plasmalogens, glycoglycerolipids, and glycosphingolipids form vesicles and bilayers.

Question: How will (1) length of hydrocarbon chain, (2) ionic strength of the aqueous medium, and (3) concentration of amphiphile affect the size of small micelles?

1. Micellar size is larger with lipids having longer hydrocarbon chains, since the hydrocarbon-volume-to-head ratio is higher and the attractive forces between the tails (van der Waals forces) are greater than the ionic repulsive forces between the head groups.
2. Increasing ionic strength of the aqueous medium will permit formation of larger micelles since the higher ionic strength will decrease the ionic repulsive forces between the head groups and thus more lipid molecules will pack together.
3. At very low concentrations of amphiphile, micelles will not form; the transition from the unaggregated to the micellar state occurs over a narrow range of concentration known as the *critical micellar concentration*.

Question: What role can cholesterol play in micellar formation?

Cholesterol does not form micelles because (1) it is not amphiphilic and (2) its flat, rigid, fused-ring structure gives a solid rather than a liquid, mobile hydrocarbon phase necessary for micellar formation. Cholesterol forms *mixed micelles* with amphiphilic lipids and will enter monolayers.

6.8 BILE ACIDS AND BILE SALTS

The *bile acids* are produced in the liver by the metabolism of cholesterol. They are di- and trihydroxylated steroids with 24 C atoms. The structure of cholic acid was seen earlier (Sec. 6.6). *Deoxycholic acid* and *chenodeoxycholic acid* are two other bile acids. In the bile acids, all the hydroxyl groups have an α orientation, while the two methyl groups are β. Thus, one side of the molecule is more polar than the other. However, the molecules are not planar but *bent* because of the cis conformation of the A and B rings.

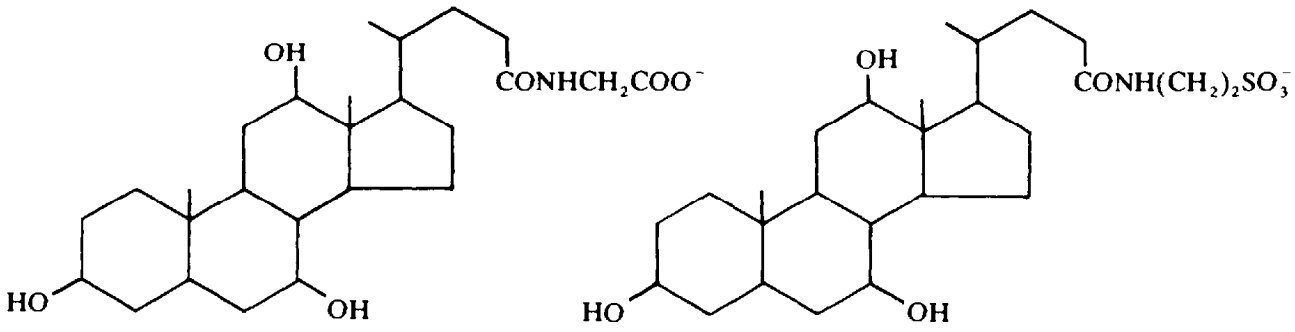
The bile acids produced by the liver accumulate in the gall bladder in the form of *bile salts*; they are bile acids in which the carboxylic acid group is conjugated with glycine or taurine.

$^{+}NH_3CH_2COO^{-}$ Glycine

$^{+}NH_3CH_2CH_2SO_3^{-}$ Taurine

EXAMPLE 6.8

The structures of the salts of glycocholic acid and taurocholic acid are:



Glycocholate Taurocholate

The bile salts have *detergent properties*, but they do not form typical micelles.

Question: Why don't the bile salts form typical micelles?

Although bile salts possess a polar head, the hydrocarbon tail is not pure hydrocarbon since there are two or three hydroxyl groups on one side of the molecule. Moreover, like cholesterol (which

also does not form typical micelles), the rigid ring system would give a tightly packed, almost solid nonpolar phase, rather than a liquid nonpolar one. Bile salts, however, do form *mixed* micelles with phospholipids.

The precise structure of aggregated bile salts is unknown, but it almost certainly involves hydrogen bonding between the hydroxyl groups in adjacent molecules. Nevertheless, the bile salts are potent detergents, and they are able to emulsify dietary lipid in the intestine, thereby making the lipid more accessible to attack by digestive enzymes. They are also required for the absorption of digested dietary lipids into the cells of the intestinal mucosa.

6.9 PLASMA LIPOPROTEINS

Blood plasma contains a number of soluble *lipoproteins*, which are classified, according to their densities, into four major types. These lipid-protein complexes function as a lipid transport system. Isolated lipids are insoluble in blood, but they are rendered soluble, and therefore transportable, by combination with specific proteins, the so-called lipoproteins. There are four basic types in human blood: (1) *chylomicrons*, (2) *very low density lipoproteins* (VLDL), (3) *low-density lipoproteins* (LDL), and (4) *high-density lipoproteins* (HDL). Their properties are summarized in Table 6.2.

The different compositions of the plasma lipoproteins give a clue to their function. Essentially, those lipoproteins rich in TAGs are synthesized by the liver (VLDL) and small intestine (chylomicrons) and deliver the neutral fat to extrahepatic tissues (particularly adipose tissue). The fat-depleted lipoproteins have a higher density, and are involved in essential cholesterol transfers.

Table 6.2. Plasma Lipoproteins

	Chylomicrons	VLDL	LDL	HDL
Density ($g\,mL^{-1}$)	<0.95	0.95–1.006	1.006–1.063	1.063–1.21
Max. diameter (nm)	500	70	25	15
% composition:*				
Protein	2	10	22	33
TAG	83	50	10	8
Phospholipid	7	18	22	29
Cholesterol and cholesterol esters	8	22	46	30

*Based on dry weight of the whole lipoprotein

Question: What is the general structure of a lipoprotein particle?

See Fig. 6-4. The polar surface of the spherical particle renders the assembly soluble in water. This structure can be considered to be a tentative one only. The amount of polar material in chylomicrons and VLDL is astonishingly small. Moreover, when lipoproteins come into contact with the membranes of the cells of target tissue, the proteins remain soluble and do not become incorporated into the membrane. This suggests that the proteins of lipoproteins have unusual properties. It is known that several species of proteins (*apoproteins*: AI, AII, B_{48}, B_{100}, CI, CII, CIII, D, and E) occur. The amino acid sequences of some of them have been determined, and they possess hydrophobic regions; i.e., they have properties suggesting that parts of their structure are compatible with hydrocarbons (e.g., TAGs and the tails of phospholipids).

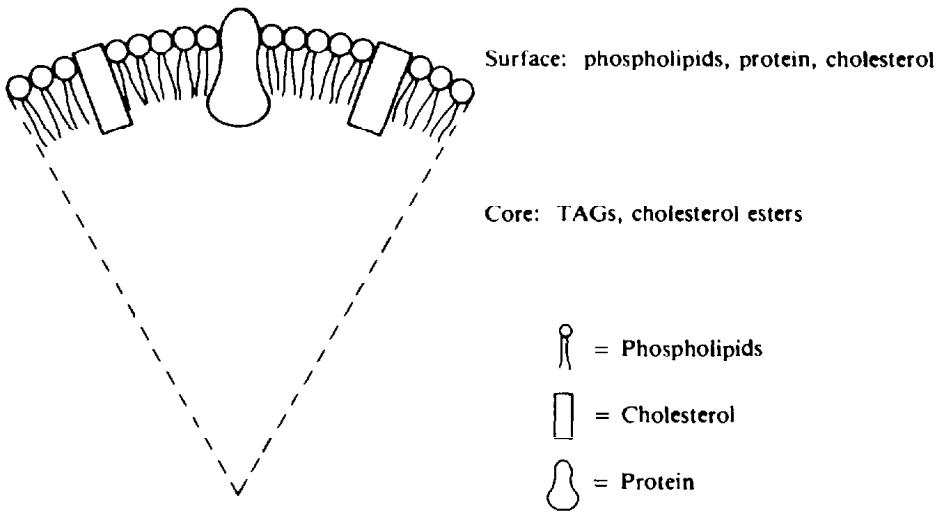


Fig. 6-4

6.10 VESICLES

When water is added to certain dry phospholipids with long hydrocarbon chains, the phospholipids swell, and when they are dispersed in more water, structures known as *liposomes* are formed. Liposomes are vesicles with multilayers of phospholipid. See Fig. 6-5. When subjected to ultrasonic vibration (*sonication*), liposomes are transformed into vesicles that have only a single bilayer of phospholipid.

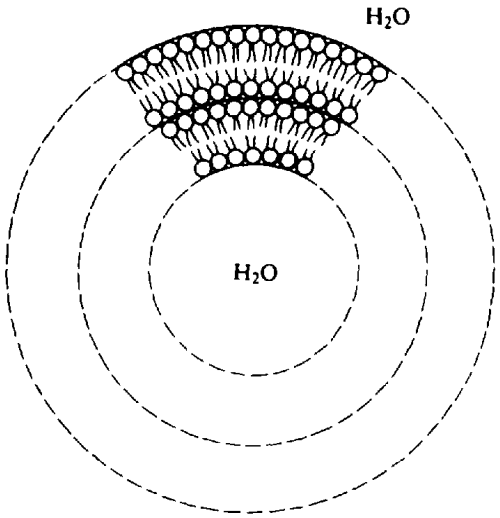


Fig. 6-5

Question: Vesicles can also be made by dialyzing (Chap. 4) a solution of phospholipid in detergent. How do vesicles form under these circumstances?

The phospholipid and detergent form mixed micelles, dominated by the detergent with its single chain of hydrocarbon; the micelles are therefore small. Dialysis lowers the concentration of the water-soluble detergent, so that the micelles become dominated by the phospholipid, which, having

two hydrocarbon chains per molecule, is more bulky. This leads to the formation of bilayers by the coalescing of micelles.

Question: Vesicles made from a mixture of phospholipids have an asymmetric distribution of the lipids between the two leaves of the bilayer. Why is this so?

The high curvature of a vesicle can make the external surface area up to about three times the internal surface area. Thus there are more than twice the number of lipid molecules in the outer *leaf* (also called *leaflet* or *layer*). The outer leaf is thicker because increasing constriction at the center of the bilayer forces the hydrocarbon chains to be fully extended. Phospholipids with large heads (e.g., PC) tend to partition into the outer leaf, where repulsions are weaker in the less constricted space on the outside of the vesicle. See Fig. 6-6.

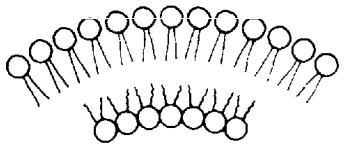


Fig. 6-6

Vesicles made by the methods described in the preceding questions are virtually impermeable to small cations and to most large polar molecules. They are slightly permeable to Cl^-, and the permeability of water is high because the solubility of water in liquid hydrocarbon is significant. When proteins are present during vesicle formation, they may be incorporated into the phospholipid bilayer. Such vesicles are known as *proteoliposomes*.

Apart from purely artificial vesicles made from phospholipids and proteins of choice, it is also possible to make vesicles from the cytoplasmic membranes of cells, usually by a sonication procedure. Artificial vesicles and vesicles derived from natural membranes have proved very useful in studying *transport* phenomena across membranes. Vesicles also occur naturally, e.g., by the budding of the Golgi apparatus in eukaryotic cells (Chap. 1).

6.11 MEMBRANES

The cytoplasm of cells is surrounded by a *plasma membrane*, and subcellular structures such as the nucleus, lysosomes, and mitochondria are delimited by membranes. The membrane of the endoplasmic reticulum in eukaryotes encloses a large intracellular space within the cytoplasm, and mitochondria have a highly folded internal membrane.

Question: What functions are served by these membranes?

They separate the cell from its environment, and they separate the different parts of the cell from each other, thus allowing certain activities to occur independently. Thus, a membrane is a physical barrier that, given the appropriate selective permeabilities, will allow the space enclosed by it to acquire and exclude useful and harmful substances, respectively, and to effect the efflux of selected compounds. Membranes also provide an environment in which chemical reactions that require nonaqueous conditions can occur.

Membranes contain lipids, proteins, and small amounts of carbohydrate. The mass ratios of these vary considerably according to the type of membrane. The carbohydrate is present as glycoglycerolipid, glycosphingolipid, and glycoprotein. The most common types of lipid found in all membranes are phosphoglycerides and phosphosphingolipid (sphingomyelin). Cholesterol is found

Table 6.3. Membrane Composition

Membrane	Component (percent of weight of membrane)			Lipid Composition* (percent of total lipid)†									
	Protein	Lipid	Carbohydrate	PA	PC	PE	PI	PS	PG	DPG	SM	GS	Chol
Human erythrocyte	49	43	8	2	19	18	1	8	—	—	18	10	25
Human myelin	18	79	3	1	10	20	1	9	—	—	8	26	26
Rat liver:													
Plasma	55	40	5	1	19	12	4	9	—	—	14	—	30
Outer mitochondria	50	47	3	1	48	22	12	2	2	3	5	—	5
Inner mitochondria	75	23	2	1	43	24	6	1	2	18	2	—	3
Escherichia coli plasma	75	25	—	tr‡	—	65	—	tr	18	12	—	—	—

* SM ≡ sphingomyelin
GS ≡ glycosphingolipid
Chol ≡ cholesterol
†Some lipid analyses do not give a total of 100% because not all lipids present are recorded here.
‡tr ≡ trace.

in plasma membranes of animals but seldom in plants. Glycosphingolipids are found in the membranes of nerve and muscle tissue. Although the type and amounts of polar heads vary widely, the hydrocarbon chains found in all membrane lipids are similar.

The composition of some membranes are shown in Table 6.3.

Question: Considering (1) the nature of the lipids found in membranes and (2) the fact that membranes are two-dimensional structures, what is the most likely arrangement of the lipid molecules in a membrane?

It is a closed bilayer as occurs with artificial vesicles (Fig. 6-3). Naturally occurring membranes are sometimes called *biomembranes* to distinguish them from the membranes present in artificial vesicles and the vesicles made experimentally from natural membranes.

Some of the proteins can be removed from membranes by agents that disrupt ionic and polar bonding (e.g., urea or high concentrations of salt solution). These are known as *extrinsic* or *peripheral* proteins. Other proteins, called *intrinsic* or *integral* proteins, can be removed only by treating the membranes with detergents or with organic solvents.

Question: What do the observations concerning extrinsic and intrinsic proteins suggest regarding the location of proteins in membranes?

If a membrane is essentially a bilayer of polar lipids, then the peripheral proteins exist on the surfaces of the bilayer and are attached via ionic and polar bonds to the polar heads of the lipids or the integral proteins. The integral proteins are deeply embedded in the bilayer and are anchored in the membrane by van der Waals bonds and hydrophobic interactions.

Question: Some of the integral proteins isolated from membranes have high molecular weights. How might these fit into the lipid bilayer?

They do this by completely spanning the bilayer (see Fig. 6-7). The parts that project on either side are polar, while the parts embedded in the bilayer consist of amino acid residues with hydrophobic

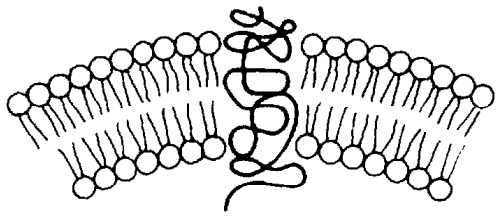


Fig. 6-7

side chains or are folded in such a way that the hydrophilic side chains project into the center of the structure away from the lipid. Integral proteins may pass through the membrane up to 12 times or more.

Although relatively few integral membrane proteins have been purified, a large number of integral proteins have now been sequenced. In these proteins it has proved possible to identify *transmembrane domains* as sequences of 20 amino acids that have a high content of hydrophobic amino acids.

So far, no protein has been found as a common constituent of all membranes (compare the almost universal existence of the lipids PC and PE), even from the same species. Thus, it seems unlikely that there is a universal structural protein in membranes. The numbers of different proteins in a membrane vary widely according to membrane type. The plasma membrane of the bacterium *Halobacterium halobium* contains only 1 protein (bacteriorhodopsin), whereas the membrane of another bacterium, *Escherichia coli*, contains about 100. The plasma membrane of the human red blood cell contains at least 17 different proteins.

The arguments developed from the beginning of this section concerning the occurrence and relationship between the lipids and proteins in membranes led S.J. Singer and G.L. Nicolson in 1972 to propose the so-called *fluid mosaic model* as a universal scheme for membrane structure (Fig. 6-8).

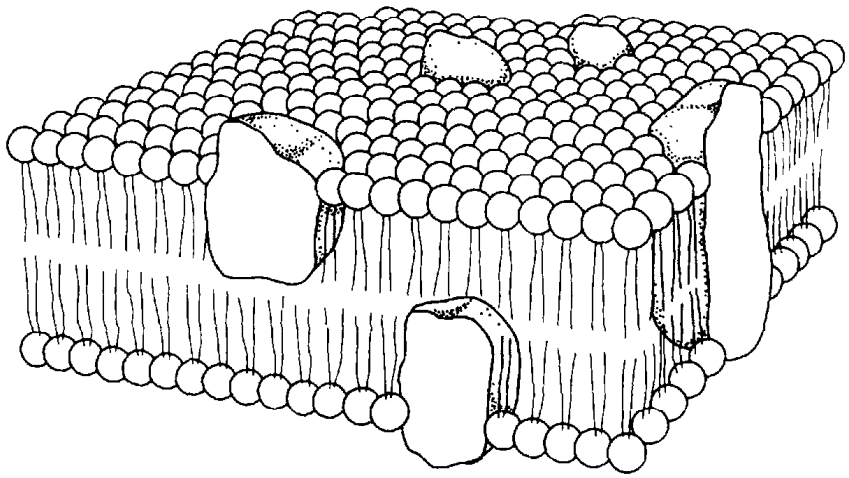


Fig. 6-8 The fluid mosaic model for membrane structure. The mosaic bilayer of polar lipids is about 5 nm thick. The proteins, including a transmembrane protein, are shown as irregular lumps.

Question: The model in Fig. 6-8 envisages a fluid, dynamic structure. What is meant by this?

The structure is not rigid. Because the hydrocarbon region is liquid, there is rapid *lateral diffusion* and *rotational motion* (about an axis perpendicular to the bilayer) of both lipid and protein components.

Question: What sort of motion is restricted in the model?

Movement of components from one leaflet to the other (*flip-flop*). For this to occur, the polar regions of a lipid or a protein must pass, at some stage, through the hydrophobic core of the bilayer, and this is a thermodynamically unfavorable occurrence.

Question: What consequences does the restriction on flip-flop of lipids have for the structure of the membrane?

Membranes often show an *asymmetric distribution* of lipid and protein components between the two leaflets of the bilayer. There is ample evidence for this. For example, PC and sphingomyelin, where it occurs, are found essentially in the outer leaflet, whereas PE and PS are found predominantly in the inner leaflet. The existence of unidirectional membrane transport indicates that the proteins involved must be asymmetrically distributed across the membrane. Moreover, the outer leaflet of the plasma membrane contains the molecules (glycoproteins and glycolipids) that identify the cell as a particular type so that it can be recognized by components in the circulation. These glycoproteins and glycolipids do not occur in the inner leaflet. Recent evidence suggests that certain phospholipids (e.g., PS) are asymmetrically distributed by the action of energy-requiring phospholipid translocases.

Question: How will temperature affect the physical properties of membranes?

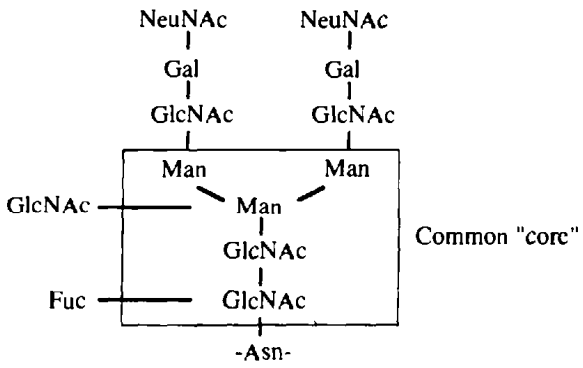
At temperatures well below that at which a membrane occurs naturally, the lipid bilayer will be nonliquid. A reversible transition, usually over a range of $\sim 10°C$, occurs as the temperature is raised, and the hydrocarbon chains become disordered as the membrane becomes liquid. The midpoint of this transition is known as the *melting point* or *transition temperature*. Its value in a particular organism is maintained a few degrees below the ambient temperature and depends on (1) the lengths of the hydrocarbon chains and their degree of unsaturation and (2) the nature of the head groups of the lipids.

The dependence of the melting point upon the nature of the head groups of the lipids suggests that interactions occur between the various head groups. This is supported by the observation that PE is much less mobile than PC in artificial membranes. With PE, hydrogen bonds can occur between $-NH_3^+$ and a phosphate on a neighboring molecule, whereas the larger group $-{}^+N(CH_3)_3$ in the PC molecule relative to $-NH_3^+$ makes it less likely that $-{}^+N(CH_3)_3$ will fit into a tightly packed array without creating a stronger repulsive force; this leads to greater mobility.

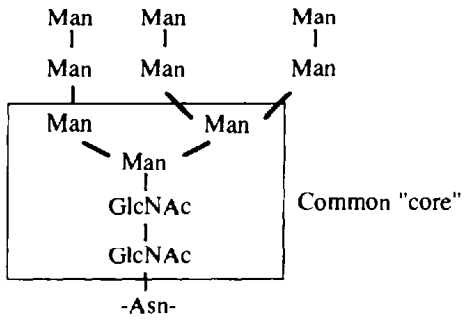
A number of integral proteins retain bound lipid molecules when isolated. It is possible, then, that, in the intact membrane, the mobility of protein and the surrounding layer of lipid is restricted by a physical association between the two. Removal of the *lipids of solvation* causes loss of structural and functional integrity of the protein.

Question: How does the presence of cholesterol affect the properties of a membrane at physiological temperature?

Cholesterol decreases the freedom of motion of the lipids surrounding it, since cholesterol is a rigid structure having a high affinity for the hydrocarbon chains of its neighbors. Moreover, it is deeply embedded, so that its hydroxyl group is level with the ester bonds of neighboring lipids and there is hydrogen bonding between the hydroxyl group and the polar heads. In general, cholesterol-rich membranes have reduced water permeability.



(*a*) Complex type



(*b*) High mannose type

Fig. 6-9 N-Linked oligosaccharides in glyoproteins: (*a*) complex type, and (*b*) high mannose type.

Membrane Glycoproteins

The recognition of cells by other cells often involves the presence of carbohydrates attached to proteins on the surface of cell membranes. Such carbohydrates are frequently attached to serine or threonine residues (Chap. 3) via *O*-glycosidic linkages, or to the side chains of asparagine residues via *N*-glycosidic linkages. The *N*-linked oligosaccharides frequently contain a common "core" of 5 residues, two of *N*-acetylglucosamine and three of mannose, as shown in Fig. 6-9. To this core may be attached additional mannose residues (the high mannose type) or various combinations of *N*-acetylglucosamine, galactose, *N*-acetylneuraminic acid and fucose (the complex type). With a number of different monosaccharide units to choose from, several possibilities of linkage, and two anomeric configurations, an enormous number of possible structures is possible.

Question: What is the point of having such diversity?

A large degree of possible diversity provides a large information content, allowing cells to develop complex mechanisms of recognition.

EXAMPLE 6.9

It is believed that senescence ("passing the use-by date") of mammalian cells or circulating glycoproteins may be recognized by the loss of *N*-acetylneuraminic acid (sialic acid) brought about by *sialylases* in vascular tissues. The exposed terminal galactose residue may be recognized and bound by an *asialoglycoprotein receptor*

of certain cells which can then internalize and degrade the cell or protein carrying the asialylated glycoprotein by means of endocytosis.

EXAMPLE 6.10

Plants and some bacteria produce proteins, known as *lectins*, that are capable of specific binding to certain classes of oligosaccharides. A lectin from jack bean, known as *concanavalin A*, recognizes and binds to nonreducing α-mannose units of oligosaccharides. Bacteria, such as *E. coli*, attach to host tissues through bacterial lectins that recognize and bind particular carbohydrate units on the surface of host cells. Recognition mechanisms of this type explain the tissue specificity of some bacterial diseases.

6.12 TRANSPORT

A membrane is primarily a mechanical barrier separating two aqueous phases and is able to function as such by virtue of the two-dimensional hydrophobic bilayer of polar lipids that inhibits the free movement of solutes from one phase to the other. Free movement across a membrane is referred to as *simple diffusion*. The rate of simple diffusion of different substances varies considerably. For example, the following substances are arranged in order of decreasing rate of simple diffusion across a typical membrane: oxygen, benzene, glycerol, glucose, aspartate, hemoglobin. The rate of simple diffusion depends on the state of the substance (gases diffuse fastest), its size (the smaller, the faster), and its polarity (the more polar, the slower). In quantitative terms and for nongases, simple diffusion is of importance only in the transport of molecules with a large hydrophobic character. A few membranes, however, permit rapid simple diffusion of a wide range of substances, including polar compounds.

Question: By what mechanism do some membranes allow nonselective entry of certain solutes?

Some membranes contain *pores*. Examples of membranes that contain pores and allow quite hydrophilic compounds, generally of M_r less than 600, to pass freely include the outer membranes of mitochondria and those of Gram-negative bacteria. The pores of the membranes are formed by proteins called *porins*.

Most membranes do not possess nonspecific pores allowing rapid diffusion of solutes, and for them, simple diffusion of solutes is a very slow process. Solutes can cross such membranes much more rapidly by *carrier-mediated transport*. The membrane components that mediate transport are proteins. They are known as *carriers* or *permeases* or *transport proteins*. The term *translocase* has sometimes been used but should be avoided since this term is also used to describe enzymes involved in protein biosynthesis.

Carriers are specific for the solute transported. A set of carriers will allow a cell to take up desirable substances rapidly and get rid of unwanted substances rapidly. By the same token, undesirable compounds can be kept out of cells, and useful ones retained, by the *absence* of suitable carriers, in which case simple diffusion would be the only way such compounds could cross membranes.

Although it is a simple matter, experimentally, to determine the *characteristics* of the movement of a particular compound across a membrane (e.g., the effect of pH and temperature on the rate), it is difficult to examine *how* a carrier operates at a molecular level.

Question: Why is determining the mechanisms of carrier transport so difficult?

For a carrier to be studied effectively it must usually be purified. But purifying it involves removing it from the membrane in which it exists naturally; i.e., the membrane must be disrupted, and as a result, the very quality by which the biological activity of the carrier can be measured is lost. This problem can be overcome by incorporating the purified carrier into artificial vesicles.

EXAMPLE 6.11

In order to measure the rate of transport of a solute into a cell (or artificial vesicle), radioactively labeled solute is used. A suspension of cells is incubated in a medium containing radioactive solute; portions of the suspension are removed at different times, and the radioactivity of the washed, filtered cells is measured. Fig. 6-10 shows a plot of the radioactivity of the washed cells at different times. The slope of the initial part of the plot gives the rate of uptake of the solute (after solute concentrations are first derived from measurements of radioactivity) corresponding to the solute concentration at time = 0, $[S]_o$.

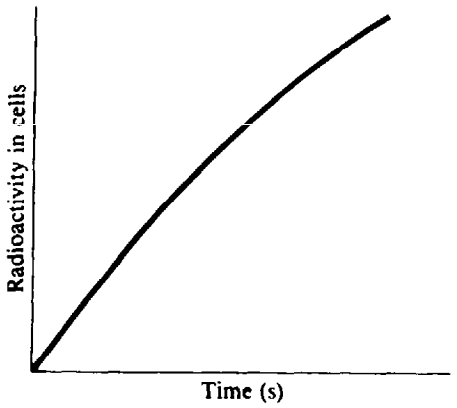


Fig. 6-10

Question: Are there any experimental artifacts likely in the method described?

Yes. Once in the cells, the solute may be metabolized. Thus, its actual concentration inside the cell will be less than that indicated by the measured radioactivity of the cells. Also, if there is solute already inside the cell, the measured rate of uptake is a *net* value since there must be some movement (by simple diffusion and carrier transport) of the solute out of the cell. (These objections do not usually apply to transport by carriers incorporated into artificial vesicles.) These problems are overcome partly by using a nonmetabolizable analog; e.g., in studying the transport of glucose, radioactive 2-deoxyglucose or α-methylglucoside could be used.

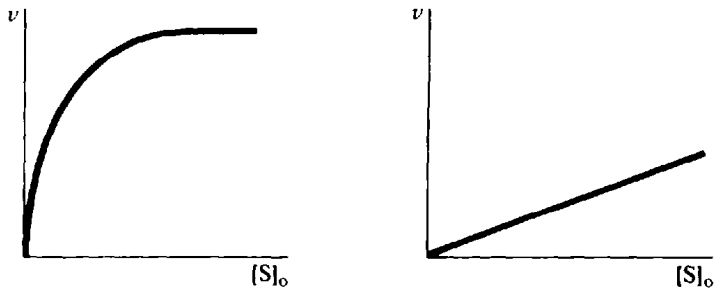
In studies on transport in bacteria, another method for overcoming the problem of solute metabolism is to use bacterial mutants which will transport the solute but which lack one or more enzymes for metabolizing the solute.

Question: What would be an objection to the use of analogs?

The precise rate of transport of the compound of interest, e.g., *glucose*, is not determined, although it can be *presumed* to be similar to that of the analog.

Question: Given that it is possible to measure v, the rate of transport of a solute into a cell, how will v vary with the concentration of solute $[S]_o$ in which the cells are suspended for (*a*) carrier-mediated transport and (*b*) simple diffusion?

For carrier-mediated transport [Fig. 6-11(*a*)] there must be a finite number of carrier molecules in the membrane. At low $[S]_o$, only some of these molecules will be bound to the solute, but at high $[S]_o$, *most* of the carrier molecules will be occupied, and there is therefore a maximal value for v (V_{max}). In simple diffusion [Fig. 6-11(*b*)], there is no carrier to *saturate*, and v is higher at high $[S]_o$ because the concentration gradient of solute across the membrane, which determines the rate of diffusion, is greater than it is at low $[S]_o$.



(*a*) Carrier-mediated transport (*b*) Simple diffusion

Fig. 6-11 Dependence of rate of transport on solute concentration.

Mathematical equations can be derived to explain the curves demonstrated in Fig. 6-11 (see Chap. 9). The curve in Fig. 6-11(*a*) is described by the equation for a rectangular hyperbola.

$$v = \frac{V_{max}[S]_o}{[S]_o + \text{a constant}} \tag{6.1}$$

The equation for the graph in Fig. 6-11(*b*) is

$$v = \frac{\text{a constant}}{l}[S]_o \tag{6.2}$$

where l is the thickness of the membrane. (In the derivation of these equations it is assumed there is no solute on one side of the membrane when a concentration of $[S]_o$ is introduced on the other side.)

Question: Suppose carrier-mediated transport *and* appreciable simple diffusion occurred simultaneously in an experiment on transport. How could this be demonstrated?

A plot of v against $[S]_o$ would not show saturation, except when simple diffusion was trivial compared with carrier-mediated transport. See Fig. 6-12. However, it is often impossible to determine v at high $[S]_o$; e.g., the solute may not be soluble enough to achieve high $[S]_o$. Mixed processes of movement of solute across a membrane can be demonstrated using a *double-reciprocal plot*. (See also Chap. 9.)

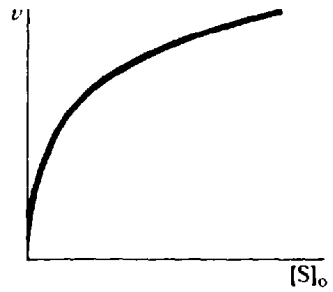


Fig. 6-12

EXAMPLE 6.12

The inverted forms of Eqs. (*6.1*) and (*6.2*) are

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{\text{a constant}}{V_{max}}\frac{1}{[S]_o} \quad \text{and} \quad \frac{1}{v} = \frac{l}{\text{a constant}}\frac{1}{[S]_o}$$

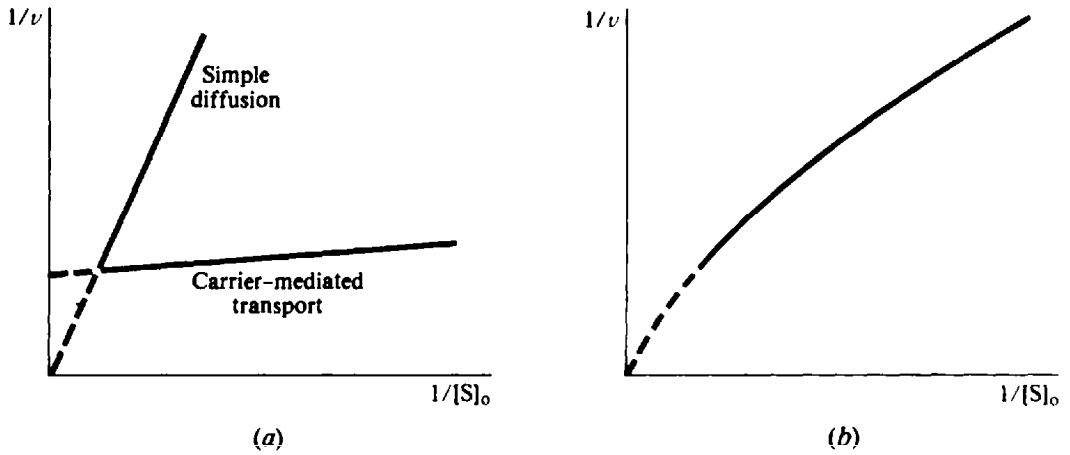


Fig. 6-13

Plots of $1/v$ against $1/[S]_0$ for both of these are straight lines [see Fig. 6-13(*a*)]. For simultaneous simple diffusion and carrier-mediated transport, the double-reciprocal plot is shown in Fig. 6-13(*b*).

An alternative method for distinguishing simple diffusion from carrier-mediated transport can be used if the solute contains a chiral atom. Simple diffusion will occur for both enantiomers and at the same rate, but carrier-mediated transport is invariably stereospecific; i.e., the carrier will recognize only one of the enantiomers.

Question: Which would be *recognized* (bound) by a carrier located in a cell membrane, D-glucose or L-glucose?

D-Glucose would be recognized by a carrier because this isomer occurs naturally. The other does not, and there would be little point in living systems evolving such a carrier.

Question: Under what conditions would the process of (*a*) simple diffusion and that of (*b*) carrier-mediated transport cease?

(*a*) Simple *net* diffusion would cease when the concentrations of solute on both sides of the membrane were the same.

(*b*) With some carrier-mediated transport systems it has been found that, as in simple diffusion, net transport ceases when the solute concentration is the same on both sides of the membrane. With other systems, the solute continues to be transported even after the solute concentrations on both sides become equal.

As the preceding question shows, there are *two* types of carrier-mediated transport: (1) *facilitated diffusion* (which allows the concentration of solute on both sides of a membrane to be equalized) and (2) *active transport* (which allows the solute to move *up*, or *against*, a concentration gradient).

EXAMPLE 6.13

Figure 6-14 shows the ways in which solutes can move across a membrane. [S] and [s] represent high and low concentrations of solute, respectively.

Simple diffusion and facilitated diffusion are spontaneous processes; solute will move down a concentration gradient (i.e., [S]→[s]) until equilibrium is attained. The free energy change, ΔG, for

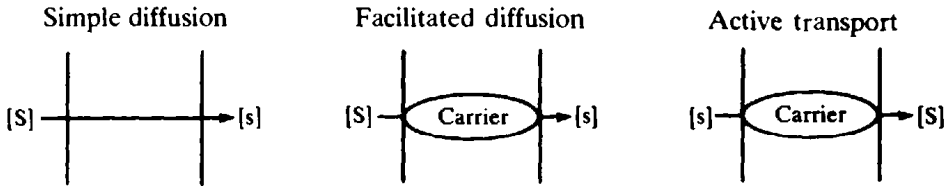


Fig. 6-14

these processes is negative (Chap. 10) because the solute becomes more randomly distributed. Thus, no input of energy is required for the process.

$$\Delta G = 2.3RT \log \frac{[s]}{[S]} \tag{6.3}$$

At equilibrium [s] = [S] and ΔG is zero.

With active transport, the solute is concentrated on one side of the membrane. This process is not spontaneous, since ΔG is positive and some form of energy must be provided to drive the carrier-mediated reaction. Hence the term *active* transport.

Question: Is the ability to concentrate a particular solute by active transport limitless?

No. If the concentration of solute builds up to high levels on one side of a membrane, solute leaks back by simple diffusion. The higher the concentration difference across the membrane, the greater will be the rate of simple diffusion. In practice, concentration ratios greater than several hundred-to-one rarely occur, and the ratios are usually very much smaller.

The solute will also diffuse back across a membrane if the membrane contains a carrier for facilitated diffusion. That is, simple diffusion and facilitated diffusion are *bidirectional*; active transport is *unidirectional*.

Question: The movement of an *ionized* solute across certain types of membranes does not always give rise to equal concentrations of the ion on both sides of the membrane, even when only simple diffusion, facilitated diffusion, or both are available. How can this be explained?

If an *electrical potential difference* exists across the membrane, then the ionized solute will be

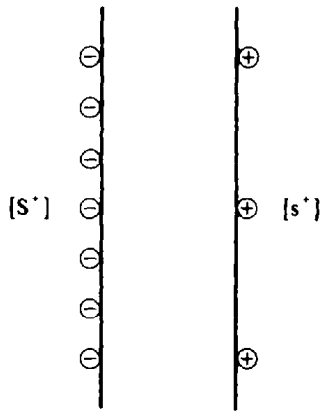


Fig. 6-15 [S^+] and [s^+] = equilibrium concentrations. Note that there are more negative than positive charges on the membrane.

unevenly distributed at equilibrium (Fig. 6-15). ΔG is dependent not only on solute concentration, but also on the potential difference $\Delta\Psi$ across the membrane (Chap. 10).

$$\Delta G = 2.3RT \log \frac{[s^+]}{[S^+]} + ZF\Delta\Psi$$

where Z is the net charge on the solute, and F is the Faraday constant.

For active transport to occur, some form of energy must be provided. In *primary active transport*, energy is provided either by the hydrolysis of ATP or by utilization of electron flow down an electron-transport chain (see Chap. 10). In *secondary active transport*, the energy is provided by ions moving down a concentration gradient coupled to the movement of a solute against its concentration gradient. The ion gradient is set up by primary active transport. A special type of active transport is *group translocation*. In this, the substance transported is modified chemically while crossing the membrane, e.g., by accepting a chemical group; energy is expended since the chemical group is transferred as a result of cleavage of a *high-energy* compound.

EXAMPLE 6.14

Figure 6-16 shows the three types of active transport. The source of energy is assumed to be ATP. The symbols [s], [i^+] and [S], [I^+] represent low and high concentrations of solute and ion, respectively. AX is a compound from which a chemical group X (e.g., a phosphate) is transferred to the solute to form SX.

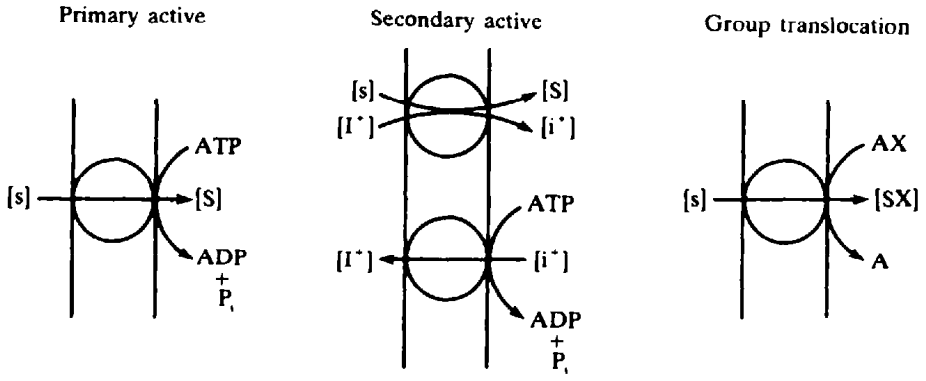


Fig. 6-16

Some examples of carrier systems are given in Table 6.4.

Table 6.4. Carrier-Mediated Transport Systems

Solute	Type of Transport	Occurrence
Glucose	Facilitated diffusion	Most animal cells
ATP out—ADP in	Facilitated diffusion	Mitochondria
H^+	Primary active (uses ATP)	Stomach epithelia
Na^+ out—K^+ in	Primary active (uses ATP)	Animal cells
Glucose	Secondary active (Na^+ cotransport)	Some animal cells
Glucose	Group translocation	Many bacteria

6.13 MOLECULAR MECHANISMS OF TRANSPORT ACROSS MEMBRANES

A protein engaged in the transport of a specific solute across a membrane must have two properties: (1) the ability to bind the solute and (2) the ability to carry out a *vectorial* process, i.e., a *directional* process which delivers the solute from one side of the membrane to the other.

Question: What are some simple means by which the vectorial process might occur?

1. The transport protein could behave as a *mobile carrier*; e.g., a large, transmembrane protein could rotate, as in Fig. 6-17(*a*), or a small protein could traverse the membrane as in Fig. 6-17(*b*).
2. The transport protein could constitute a pore or channel as in Fig. 6-17(*c*).

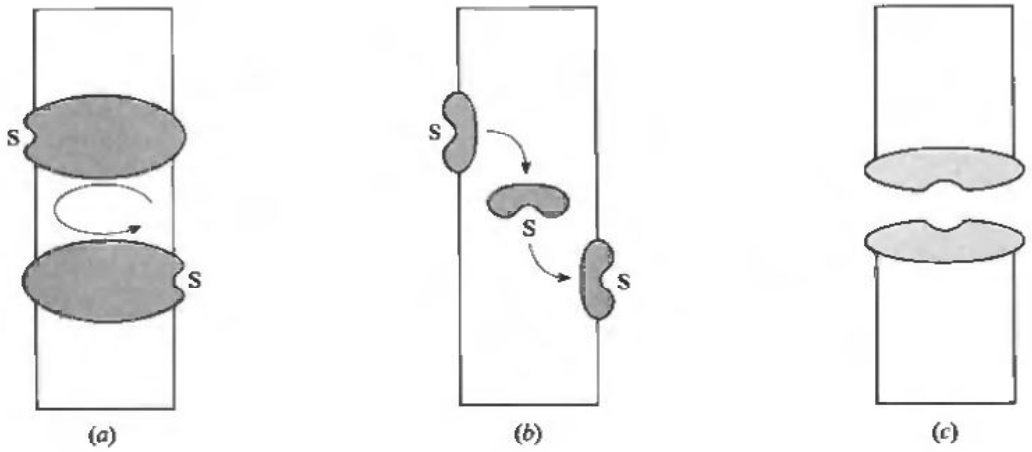


Fig. 6-17

There are difficulties with both concepts presented in the above question. Movement of polar segments of a molecule through the nonpolar part of a lipid bilayer is a rare event, and the mobile carrier concept is not favored on thermodynamic grounds. Although the existence of *nonspecific* channels in some membranes is known, the concept of a pore can be accepted only if the pore has a *gate* enabling the transport protein to be specific and to permit movement of the solute in the direction demanded by the particular type of transport. The pore concept can be modified by proposing that the transport protein is an oligomeric protein in which the spaces between subunits constitute a water-filled channel closed by contact of the subunits (Fig. 6-18). Binding of solute to the protein triggers a conformational change altering the relative positions of the subunits and so opens the channel. This would be *facilitated* diffusion. If metabolic energy is expended to cause the conformational change, rather than simply binding of the solute, then this would be *active transport*. So far, experimental studies on the structures of transport systems have provided support for the pore

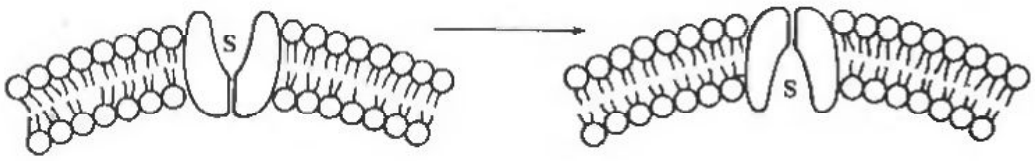


Fig. 6-18 Conformational change in a pore protein allowing transport of the solute.

concept in the case of ion channels, and for the mobile carrier hypothesis in the case of certain transport ATPases.

Artificial systems permitting the transport of cations have been inserted into natural membranes and into artificial vesicles. These are known as *ionophores*, and there are two types, exemplified by *gramicidin A* and *valinomycin*.

1. Gramicidin A is an antibiotic composed of 15 amino acids:

   CHO—NH—L-Val—Gly—L-Ala—D-Leu—L-Ala—D-Val—L-Val—D-Val—L-Trp—D-Leu—L-Trp—D-Leu—L-Trp—D-Leu—L-Trp—$CONH(CH_2)OH$

   Because of its hydrophobic nature, gramicidin A penetrates a lipid bilayer: two molecules form a head-to-tail dimer, which can twist to form a hollow helix with six turns and a total length of 3 nm, about the distance across the hydrophobic region of a lipid bilayer. The helix is stabilized by intramolecular hydrogen bonds between peptide links in the polypeptide chain. The hydrophobic side chains are on the outside of the helix, leaving a hydrophilic aqueous channel down the center of the helix through which cations can pass. This type of ionophore is nonspecific, and the cations do not bind strongly to the gramicidin A.

2. Valinomycin is an ionophore but with different properties from gramicidin A: (*a*) it specifically transports K^+, and no other ion, when inserted into membranes or vesicles; (*b*) it can transport K^+ only above the phase transition temperature of the membrane, whereas cation transport by gramicidin A is insensitive to temperature.

   Valinomycin has a cyclic structure resembling a peptide chain containing 12 amino acid residues; basically it is a four-unit structure that is repeated three times:

   $$\text{NH—(L-Val—D-HyV—D-Val—L-Lac)}_3\text{CO}$$

   where D-HyV is D-hydroxyisovalerate and L-Lac is L-lactate. The structure is stabilized by successive amide and ester linkages.

Question: What would be a possible mechanism of K^+ transport by valinomycin?

Since valinomycin transports K^+ specifically, it must bind the ion rather than provide a channel, as does gramicidin. Since transport can occur only when the membrane is fluid and since valinomycin is a small molecule, this suggests that the valinomycin-K^+ complex must move through the membrane; i.e., it behaves as a *mobile carrier*. The hydrocarbon side chains of the cyclic structure can be imagined on the outside of the structure, thus making the latter compatible with the hydrocarbon portion of the lipid bilayer. The inside contains the 12 carbonyl groups of the ester and amide bonds, six of which coordinate the K^+ in the space at the center of the structure (Fig. 6-19).

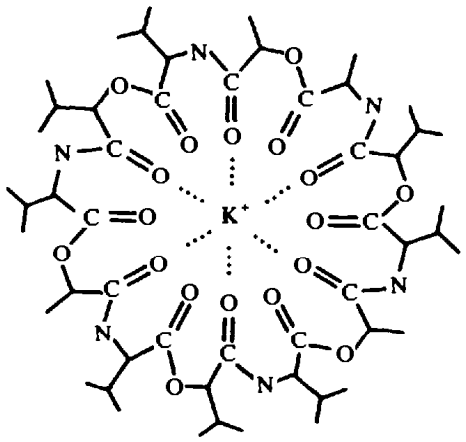


Fig. 6-19 Binding of K^+ by valinomycin.

Table 6.5. Plasma Membrane Receptors

Receptor Type	Examples	Natural Ligand	Functions
Ligand-gated ion channels	Nicotinic (Na^+, Ca^{2+} channel)	Acetylcholine	Activation of neuromuscular junction
	GABA (A) (Cl^- channel)	γ-Aminobutyrate	Inhibition of nerve transmission in brain
	Excitatory amino acid (Na^+, Ca^{2+} channel)	Glutamate	Activation of nerve transmission in brain
	P_{2X} purine nucleotide receptors (Na^+, Ca^{2+} channel)	ATP	Activation of nerve transmission in brain
G-protein coupled receptors (a GTP-binding protein links the receptor to the activation of intracellular enzymes)	Muscarinic receptors (five subtypes)	Acetylcholine	Slow heart rate; release pancreatic digestive enzymes
	Adrenergic receptors (three α1 subtypes; three α2 subtypes; three β subtypes)	Epinephrine, norepinephrine	Increase blood pressure (α1); increase heart rate (β1); increase air flow into and out of lungs (β2)
	Angiotensin receptor	Angiotensin II	Increase blood pressure (by constricting small arteries)
	Leukotriene B_4 receptor	Leukotriene B_4	Neutrophils: promote intracellular killing of bacteria
	P_{2U} purine nucleotide receptor	ATP, UTP	Endothelial-dependent relaxation of vascular smooth muscle
Growth factor receptors (These receptors fall into two main groups: receptor tyrosine kinases[1] and receptors that associate with cytoplasmic tyrosine kinases[2])	Insulin[1]	Insulin	Lower blood glucose; promote glucose storage as glycogen in liver and skeletal muscle and as fat in fat cells
	Epidermal growth factor[1] (EGF)	EGF	Promote differentiation of cells that line the gut
	Interferon (INF)[2] (several subtypes)	INF α,β,γ	Promote differentiation and proliferation of lymphocytes
	Growth hormone[2]	Growth hormone	Promote growth of diverse tissues

6.14 SIGNALING

The plasma membrane is an essential barrier between the intracellular and extracellular fluids. Signaling mechanisms have evolved so that cells may respond to chemical messengers that derive from nerve terminals (neurocrine transmission), endocrine glands (endocrine transmission) or neighboring cells (paracrine transmission).

Two basic forms of signaling are recognized.

1. An amphiphilic chemical messenger diffuses across the plasma membrane. Inside the cell the messenger molecules interact with specific receptors (binding proteins). In the ligand-bound state these receptors regulate the transcription of specific messenger RNAs and thereby control the expression of key enzymes, receptors or transporters.
2. The chemical messenger is unable to cross the plasma membrane because of low lipid solubility. The messenger binds to specific plasma membrane receptors (integral membrane proteins). In the ligand-bound state the receptors activate either (1) an intrinsic ion channel (*ionotropic receptors*) or (2) a defined sequence of enzymes (*metabotropic receptors*). Table 6.5 shows examples of plasma membrane receptors.

Nuclear Receptors

These receptors are chiefly responsible for the physiological effects of steroid hormones such as cortisol as well as thyroid hormone and vitamin A. They are proteins that share a common basic structure consisting of a ligand binding domain and a DNA binding domain (comprised of zinc finger motifs). They operate as ligand-responsive transcription factors (see Chap. 17 for further discussion).

EXAMPLE 6.15

The glucocorticoid receptor is a nuclear protein that has three functionally important domains: a domain at the C terminus that binds steroid hormone (ligand), a DNA-binding domain composed of two zinc finger motifs and an N-terminal domain involved in the regulation of gene transcription (Fig. 6-20).

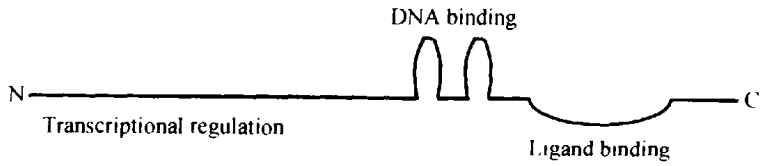


Fig. 6-20 The structure of the glucocorticoid receptor.

Plasma Membrane Receptors

Ionotropic Receptors (Ligand-Gated Ion Channels)

The chemical messenger alters the conformational state of an integral membrane protein that operates as a channel for the movement of ions (e.g., Na^+, K^+, Ca^{2+}, Cl^-) across the plasma membrane. Excitatory neurotransmitters such as glutamate, acetylcholine and ATP activate ligand-gated ion channels that promote the entry of Na^+ and Ca^{2+} ions to depolarize (activate) neurones. The inhibitory neurotransmitter γ-aminobutyrate (GABA), on the other hand, promotes the entry of Cl^- ions that hyperpolarize (deactivate) neurons.

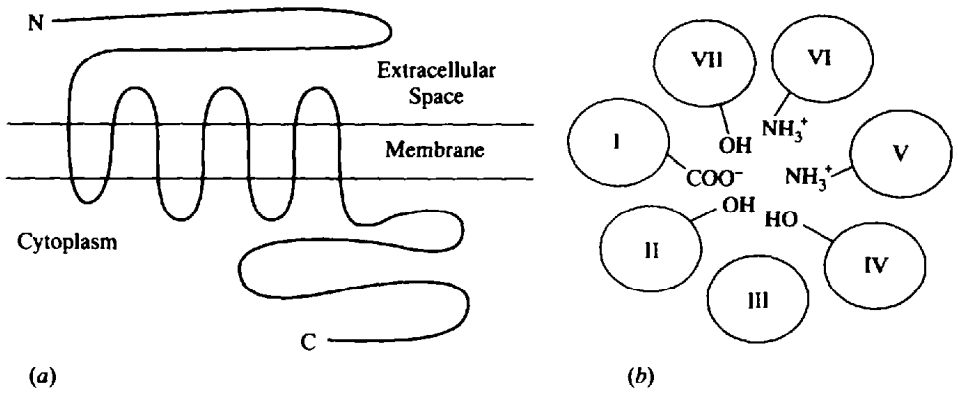


Fig. 6-21 Schematic views of G-protein coupled receptors. (*a*) Cross-sectional view of a G-protein coupled receptor demonstrating the N terminus in the extracellular space, seven transmembrane domains and the C terminus in the cytoplasm. (*b*) A view of the transmembrane domains of a G-protein coupled receptor from above demonstrating how amino acid side chains can form a precise ligand-binding site. In this example the ligand-binding site is shown to be formed by a side chain carboxylate group, two side chain amino groups and three side chain hydroxyls.

Metabotropic Receptors (G-Protein Coupled Receptors, Growth Factor Receptors)

G-protein coupled receptors respond to an astonishing variety of activators including short peptides, proteins, biogenic amines, nucleotides, lipids and even photons of light! They are single subunit integral membrane proteins with a common seven-transmembrane domain structure in the form of a so-called "helical wheel" [Fig. 6-21(*b*)].

The ligand binding site is formed at or near the surface of the membrane by amino acid side chains that project into a central cavity formed by the transmembrane domains [Fig. 6-21(*b*)]. In the ligand-bound state, these receptors activate cytosolic guanine nucleotide-binding proteins (i.e., G-proteins) that, in turn, regulate the activity of one or more key enzymes. G-proteins consist of three nonidentical subunits ($\alpha\beta\gamma$). The α subunit binds one molecule of either GDP or GTP. The GDP-bound form is inactive. Occupation of the ligand-binding site on the receptor initiates a guanine nucleotide-exchange reaction whereby GDP is displaced by GTP (Fig. 6-22). The activated α-subunit then separates from its associated $\beta\gamma$ subunits. This is a key event in the *signal transduction* process because the liberated α subunits and $\beta\gamma$ subunits bind to and activate enzymes downstream of the receptor. The α subunit also has enzyme activity. Its intrinsic GTPase limits the duration of its own activation by converting GTP to GDP in its guanine nucleotide-binding site.

EXAMPLE 6.16

Cholera toxin interferes with the function of a key G-protein in the small intestine by deactivating its intrinsic GTPase activity. The consequence is uncontrolled activation of the signal transduction pathway which produces the second messenger 3′,5′-cyclic AMP leading to a life-threatening diarrhea following the activation of ion and water secretion.

Question: How many G-protein coupled receptors are there?

The answer to this question is not known. However, homology-based molecular cloning techniques have led to the identification of literally hundreds of these receptors each of which conforms to the seven transmembrane domain pattern. The function of many of these receptors and their ligands are unknown; such receptors are referred to as "orphan" receptors. Similarly, there is a considerable diversity of G-proteins. In particular, over 30 different α-subunit types have been described.

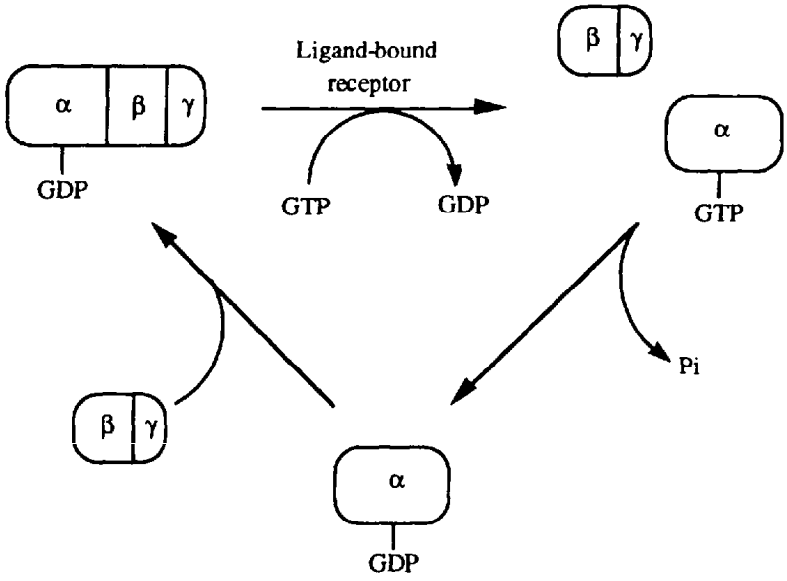


Fig. 6-22 Cycle of activation and deactivation of heterotrimeric G-proteins. The ligand-bound receptor acts as a guanine nucleotide-exchange factor that replaces GDP with GTP on the α subunit. The α-GTP and $\beta\gamma$-GTP subunits then separate to activate downstream target enzymes. The cycle is completed by the action of intrinsic GTPase activity in the α subunit. Free α-GDP subunits readily reassociate with free $\beta\gamma$ subunits to form inactive $\alpha\beta\gamma$ heterotrimers. G-protein activation represents a site of *amplification* in signaling pathways: for every ligand bound receptor molecule in the membrane, it is estimated that hundreds or thousands of G-protein molecules are activated.

Regulation of Enzymes by G-Proteins

G-proteins activate enzymes that are loosely associated with the cytoplasmic surface of the membrane or are present free in the cytoplasm. They act primarily on two forms of substrate, (1) cytoplasmic nucleotides or (2) glycerol phospholipids of the inner leaflet of the plasma membrane.

EXAMPLE 6.17

Nucleotide-Metabolizing Enzymes that are Regulated by G-Protein Coupled Receptors:

Adenylate cyclase, which converts ATP to adenosine 3′,5′-cyclic monophosphate (cAMP) is activated by receptors that activate the G-protein known as G_s (for stimulatory G-protein) and inhibited by receptors that activate the G-protein known as G_i (for inhibitory G-protein). Elevation of cAMP levels underlie the release of glucose from glycogen that is induced by epinephrine and glucagon (see Chap. 11).

cGMP phosphodiesterase, which converts guanosine 3′,5′-cyclic monophosphate (cGMP) to GMP is activated by the G-protein transducin (G_t) in response to activation of photoreceptors in the plasma membranes of retinal rod cells. Light-induced suppression of cGMP levels in this way closes Na^+-permeable channels in the rod cell membrane, resulting in membrane hyperpolarization and a decrease in nerve transmission from the retina. The brain interprets this suppression of nerve transmission as a burst of light.

G-Protein Regulated Phospholipases

Glycerol phospholipids are susceptible to enzyme-mediated attack by phospholipases at a variety of positions. Phospholipases are, in some cases, highly specific for particular water-soluble head-groups: e.g., PI-specific phospholipase C. In other cases, however, the phospholipase involved may attack phospholipids from a variety of classes: e.g., phospholipase A_2. Certain isoforms of phospholipases A_2, C and D are activated by G-proteins

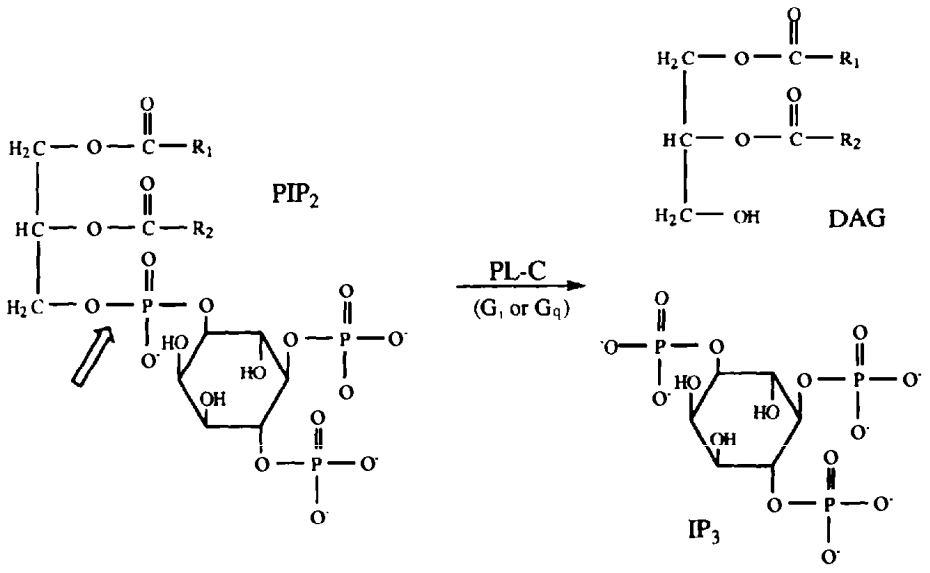


Fig. 6-23 Hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP_2) by phospholipase C. Phospholipase C (PL-C) is activated by G-protein α- and/or $\beta\gamma$-subunits. The hydrolysis site is indicated by the open arrow. Two products are obtained: 1,2-diacylglycerol (DAG) which activates protein kinase C and inositol 1,4,5-trisphosphate (IP_3) which induces Ca^{2+} release from intracellular stores.

(G_i, G_o, G_q). As a result, the inner leaflet of the plasma membrane is the source of a variety of chemical mediators that are released as a consequence of receptor activation. These mediators include: inositol 1,4,5-trisphosphate (IP_3) and 1,2-diacylglycerol (DAG), which are both products of PI-specific phospholipase C (Fig. 6-23); and arachidonic acid, which is an unsaturated fatty acid product of phospholipase A_2 and phosphatidate (a product of phospholipase D). IP_3 induces the release of Ca^{2+} ions from intracellular endoplasmic reticulum stores. DAG is a known activator of a lipid-dependent serine/threonine protein kinase (protein kinase C).

Receptors for Growth Factors

Certain growth factors such as epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) that regulate cell metabolism and gene transcription activate receptors that have intrinsic protein kinase activity in their cytoplasmic domains (Fig. 6-24). This intrinsic protein kinase activity specifically phosphorylates tyrosine residues in proteins. Receptor tyrosine kinase activity is quiescent in the absence of the ligand but is activated upon ligand binding on the extracellular surface, resulting in the phosphorylation of tyrosine residues in (1) the cytoplasmic domains of the receptors themselves (autophosphorylation) and (2) key cytoplasmic proteins. Some receptors of this type exist as single subunits that dimerize upon activation (e.g., the epidermal growth factor receptor; see below). Others consist of multiple subunits that are stabilized by disulfide bonds (e.g., the insulin receptor). Phosphorylation of the receptor's cytosolic domains creates templates for the assembly of a protein apparatus that activates a key monomeric G-protein known as *ras*. The activation of ras precedes the activation of a cascade of serine- and threonine-specific protein kinases that influence gene transcription and cytoskeletal rearrangement. These events underlie cell differentiation and proliferation.

Question: What is the nature of protein phosphorylation reactions?

Phosphorylation is a commonly used device in nature for regulating the activity of key enzymes. Protein kinases are ATP phosphotransferases that act on protein or peptide substrates. In general, protein kinases fall into two groups: those that phosphorylate serine or threonine residues (e.g.,

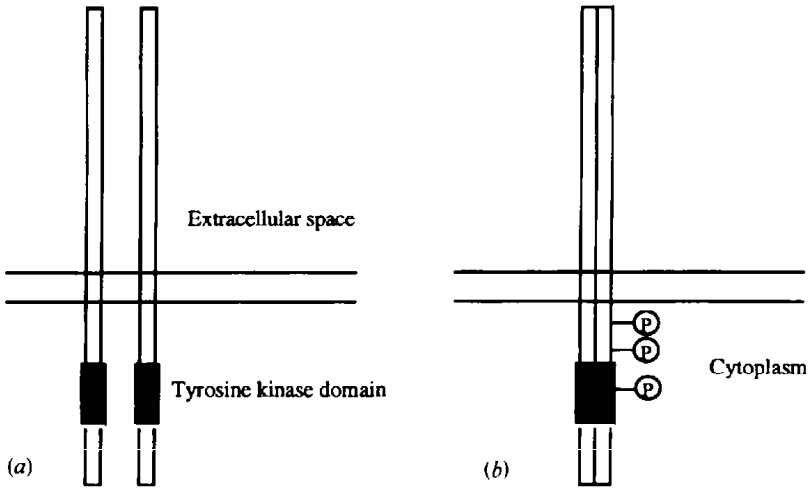


Fig. 6-24 Schematic representation of the epidermal growth factor (EGF) receptor. The receptor is an integral membrane protein with a single transmembrane domain. The ligand binding site is in the extracellular domain and there is a tyrosine kinase domain near the C terminus in the cytoplasm. (*a*) At rest the receptor exists as single subunits. (*b*) Upon binding EGF, the receptor forms dimers stabilized by noncovalent associations. After dimerization the activated tyrosine kinase phosphorylates tyrosine residues in the cytoplasmic domain prior to the recruitment of further proteins to bind to the receptor. The formation of a protein assembly on the cytoplasmic domain is necessary for activation of enzymes that regulate cell metabolism and gene transcription.

cAMP-dependent protein kinases) and those that phosphorylate tyrosine residues (e.g., the EGF or insulin receptors). In the case of serine/threonine kinases the terminal phosphate of ATP is transferred to serine and threonine hydroxyls. In the case of tyrosine kinases, the terminal phosphate of ATP is transferred to specific tyrosyl hydroxyl groups.

Question: How do proteins bind to phosphotyrosyl residues on receptor tyrosine kinases?

Specific peptide sequences known as *src* homology-2 (SH-2) domains bind to phosphotyrosyl-rich sequences. This binding reaction permits the interaction of key proteins with activated receptor tyrosine kinases and the assembly of the *ras* activating complex.

Other growth factors signal by a family of single transmembrane domain receptors that activate intracellular tyrosine kinases but do not themselves express a tyrosine kinase activity in their cytoplasmic domains. Receptors for several cytokines (peptide regulators of lymphoid and myeloid cell function) and several hormones, e.g., growth hormone, prolactin and erythropoietin, belong to this family. Upon binding of the ligand, cytoplasmic tyrosine kinases of the JAK (Janus kinase) family associate with the cytoplasmic domain of the receptor and are activated. JAK tyrosine kinases regulate key cytoplasmic enzymes and the STAT family of transcription factors.

The Role of Intracellular Mediators in Signal Transduction Pathways

Plasma membrane receptors are critical to the process of transmitting the information carried by the ligand or "first" messenger (neurotransmitter, hormone, etc.) from the extracellular space to the cell interior. However, cell signaling requires information flow inside the cell too. The molecules that carry this information to modulate the activity of ion channels, enzymes or receptors are sometimes referred to as "second" messengers. However, due to the complexity of signal transduction

cascades in which the levels of one "second" messenger can regulate the levels of another "second" messenger, the term "intracellular mediator" is probably more appropriate.

EXAMPLE 6.18

Intracellular mediators for growth factor receptors include tyrosine kinases (e.g., receptor tyrosine kinases or JAK kinases), key binding proteins, phospholipid-modifying enzymes and their products, monomeric G-proteins (e.g., ras) and downstream serine/threonine protein kinases (such as MAP kinase) which regulate the cytoskeleton and gene transcription. The activation of ras by guanine nucleotide exchange is a key event in the activation of the downstream signaling pathway for many growth factors.

Intracellular mediators for G-protein coupled receptors include cAMP, cGMP, Ca^{2+} ions, IP_3, DAG, arachidonate and phosphatidate. These mediators regulate cell metabolism by modulating the activity of ion channels (e.g., cAMP, cGMP, Ca^{2+} ions, arachidonate), serine/threonine kinases (e.g., cAMP, cGMP, Ca^{2+} ions, DAG, phosphatidate, arachidonate) or intracellular transporters (e.g., IP_3). Intracellular mediators regulate, as well as participate in, signal transduction pathways. For example, Ca^{2+} ions in concert with the Ca^{2+} binding protein calmodulin activate the breakdown of cAMP and cGMP via the intracellular enzyme phosphodiesterase (PDE). Similarly, receptor-activated protein kinases act to deactivate (down-regulate) receptors. Phosphorylation of the β-adrenergic receptor by β-adrenergic receptor kinase (BARK), for example, reduces the ability of the β-adrenergic receptor to respond to epinephrine.

Question: How do calcium ions act as intracellular mediators?

The concentration of free Ca^{2+} ions in the cytoplasm is maintained at very low levels (about 0.1 μM) under resting conditions. This is about 1/10,000 of the level observed in plasma (1.1 mM). Cytoplasmic free Ca^{2+} concentrations may increase 10- or 100-fold, however, following the activation of certain receptors. The release of IP_3 from membrane phospholipids in response to the activation of certain G-protein coupled receptors is one mechanism for increasing $[Ca^{2+}]_i$; IP_3 empties intracellular vesicular Ca^{2+} stores. The activation of certain ionotropic receptors also elevates $[Ca^{2+}]_i$ by stimulating Ca^{2+} influx from the exterior. Once elevated, Ca^{2+} activates key intracellular enzymes either directly, e.g., protein kinase C, or indirectly after binding to the high affinity Ca^{2+} binding protein calmodulin (16,700 kDa). Ca^{2+}-calmodulin activates key enzymes including specific protein kinases. Ca^{2+} levels can be restored to normal in one of several ways including the action of (1) a transport ATPase in the plasma membrane which pumps Ca^{2+} ions out of the cell or (2) a transport ATPase which accumulates Ca^{2+} ions into the interior of the endoplasmic reticulum.

Question: How does cAMP activate its protein kinase?

Cyclic AMP protein kinases (PK-A) have an $\alpha_2\beta_2$ subunit structure (Fig. 6-25). The catalytic activity resides in the α subunits which are inactive in the tetrameric $\alpha_2\beta_2$ complex. Each β subunit

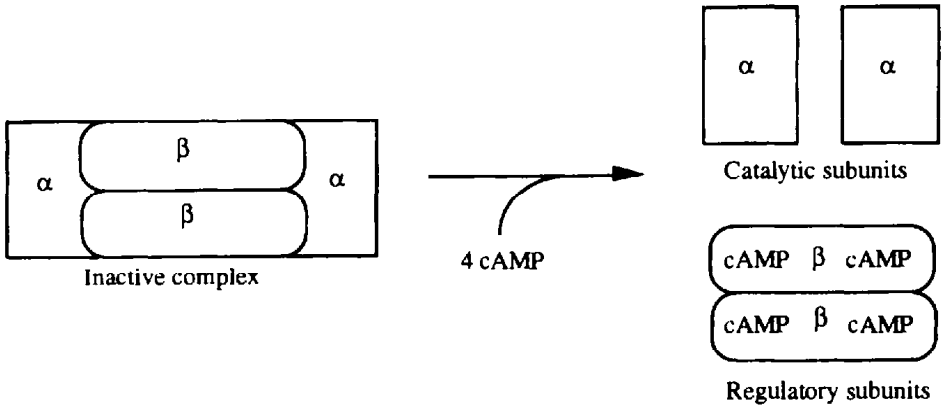


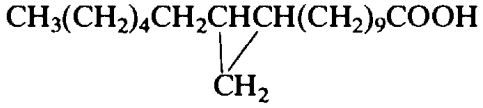
Fig. 6-25 Activation of cAMP protein kinase by cAMP.

has two cAMP binding sites. In the presence of cAMP two free α subunits are released and four molecules of cAMP bind to the remaining β_2 homodimer to form $\beta_2(\text{cAMP})_4$. The free α subunits of PK-A phosphorylate protein targets in the presence of MgATP.

Solved Problems

CLASSES OF LIPIDS

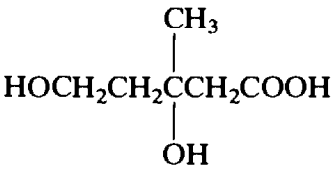
6.1. Will the melting point of lactobacillic acid ($C_{19}H_{36}O_2$, a *cyclopropane fatty acid*) be higher or lower than that of the linear, saturated fatty acid of the same chain length?



SOLUTION

The melting point will be lower; the cyclopropane group decreases the tendency of an array of molecules to pack regularly. The melting points of lactobacillic acid and the 19:0 fatty acid are 28°C and 69°C, respectively.

6.2. Mevalonic acid, radioactively labeled at the α-carbon atom, was fed to an organism that synthesizes cholesterol. Which atoms in the cholesterol will be labeled?



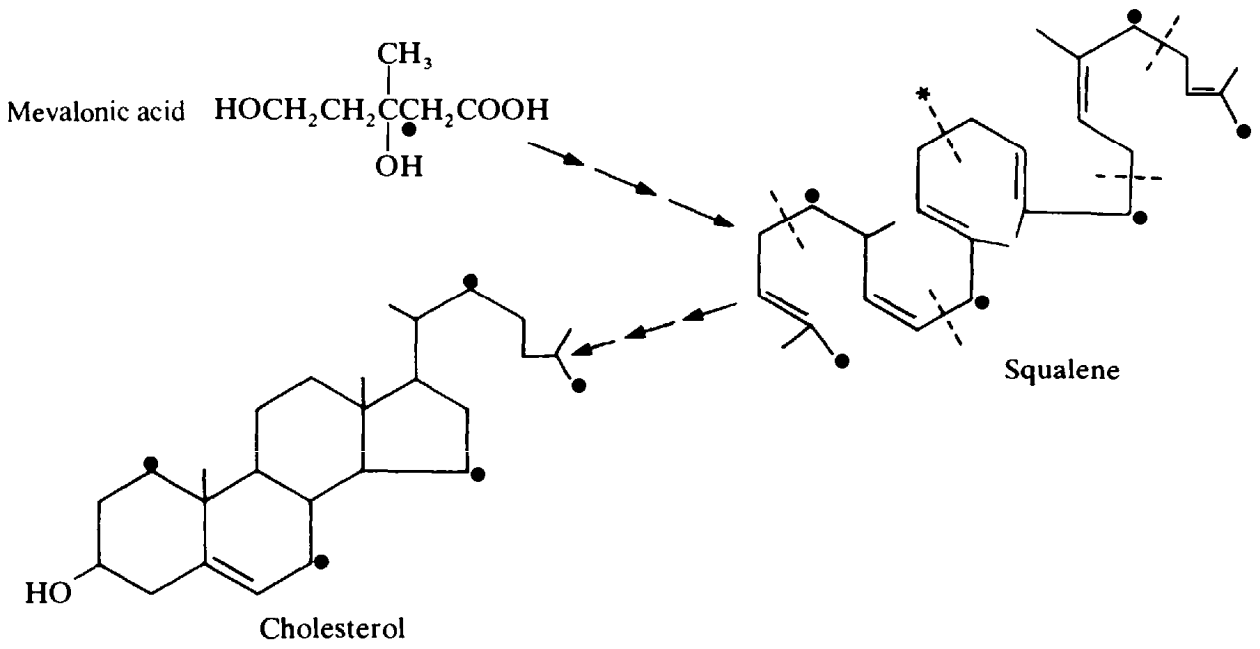
Mevalonic acid

SOLUTION

Sterols are synthesized in nature from squalene and, therefore, ultimately from isoprene. Mevalonic acid is the immediate precursor of the isoprene unit, and the carboxylic acid group is lost as carbon dioxide when two mevalonic acid molecules combine head to tail. Thus, if the α carbon of mevalonic acid is labeled, then this carbon is always adjacent to the carbon bearing a side-chain methyl group. Examination of the way in which six isoprene units are linked in squalene (Example 6.2) shows that they are not all linked head to tail; there is a point of symmetry in the structure of squalene (marked * in the structure below). At this point a set of three isoprene units, linked head to tail, is joined *head-to-head* to a similar set of three isoprene units, to give the labeling pattern shown.

BEHAVIOR OF LIPIDS IN WATER

6.3. Will phospholipids with short hydrocarbon chains form bilayers?



SOLUTION

No. They form small micelles (see the fifth question in Sec. 6.7). The hydrocarbon chains need to contain at least six carbon atoms for bilayers to form.

6.4. Explain why acylate ions have detergent properties.

SOLUTION

Acylate ions are amphiphilic, and the hydrocarbon chains are able to penetrate fatty (hydrophobic) particles, leaving the surface of the particle ionic. (See Fig. 6-26.) Thus, the particle behaves as a micelle and is readily soluble in water. The sodium and potassium salts of fatty acids are *soaps*. Soaps have poor detergent properties in hard water because the calcium present in such water causes the micelles to aggregate and precipitate. The divalent calcium ion can act as a bridge between two micelles, but since a micelle is polyvalent, a small amount of calcium relative to the amount of the soap can cause all the micelles to aggregate.

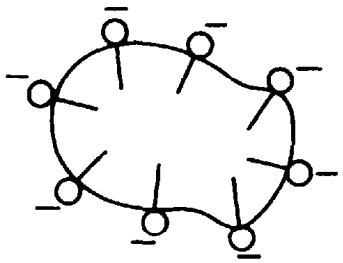


Fig. 6-26

MEMBRANES

6.5. Which would be better for solubilizing the integral proteins present in membranes, ionic or nonionic detergents?

SOLUTION

Ionic detergents alter the conformation of the hydrophobic portions of integral proteins, whereas nonionic detergents dissolve membranes and form mixed micelles of detergent-lipid-protein in which the proteins' conformations are unchanged. Thus nonionic detergents are preferred. However, removal of lipid and detergent from the mixed micelles causes conformational changes and loss of biological activity of the protein. This loss of activity arises because integral proteins need to be surrounded by *lipids of solvation* to be effective.

6.6. What strategies do bacteria adopt to restore their membrane fluidity if they are suddenly transferred from an environment at 25°C to one at 35°C?

SOLUTION

They can incorporate into their phospholipids fatty acids that are (1) longer, (2) more saturated, or (3) less branched than the originals.

TRANSPORT

6.7. To what use could L-glucose be put in studying the transport of D-glucose into cells?

SOLUTION

L-Glucose can be used for distinguishing between simple diffusion and carrier-mediated transport. L-Glucose would be transported by simple diffusion only. The difference in the rate of uptake of D-glucose and L-glucose would represent the true rate of carrier-mediated transport of D-glucose because glucose transporters are stereo-selective for D-glucose.

6.8. To each of six tubes containing buffer was added 50 μL of a suspension of rat liver mitochondria that contained 20 mg mL^{-1} of protein. Radioactively labeled pyruvate (0.07 mCi $mmol^{-1}$) was added to each tube. α-Cyano-3-hydroxycinnamate (an inhibitor of transport) was added to each tube in turn at intervals of 5 s. Each solution was filtered immediately after the addition of inhibitor, and the radioactivities of the filters (which retained the mitochondria) were found to be 0.17, 0.35, 0.51, 0.66, 0.82, and 0.96 nCi for the filters obtained at 5, 10, 15, 20, 25, and 30 s, respectively. What is the rate of uptake of pyruvate in nmol min^{-1} mg^{-1} of protein?

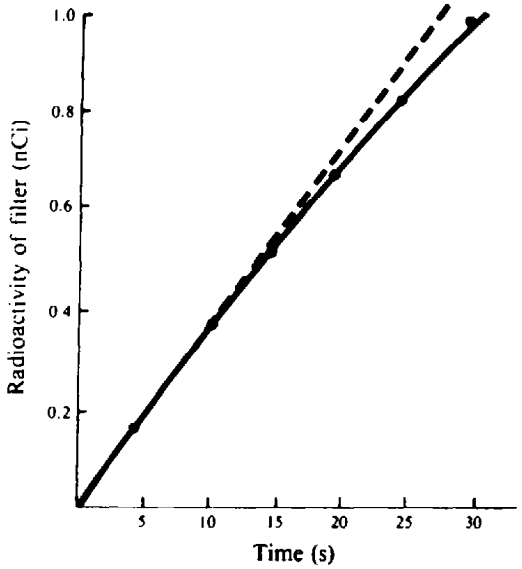


Fig. 6-27

SOLUTION

The radioactivity of the filters is a measure of the pyruvate present in the mitochondria at the time of their isolation. A graph (Fig. 6-27) of the radioactivity of the filters versus the length of time of incubation of the mitochondria with pyruvate gives a curve, but the *initial* rate of uptake of pyruvate, obtained from the slope of the tangent to the curve at the origin, is 2.1 nCi min^{-1}. This represents a rate of 30 nmol of pyruvate min^{-1} since 1 mmol of pyruvate has 0.07 mCi of radioactivity. Each incubation mixture contains 50 μL of mitochondrial suspension and therefore contains 1 mg of protein. The rate of uptake of pyruvate is therefore 30 nmol min^{-1} mg^{-1} of protein.

6.9. Suppose two independent carriers existed for the transport of a solute into a cell and simple diffusion was negligible. How could the existence of *two* carriers be demonstrated?

SOLUTION

The existence of two carriers can be demonstrated by measuring the rate of uptake (v) of the solute at different concentrations of solute ($[S]_o$). A double-reciprocal plot would have the appearance shown in Fig. 6-28. Also see Chap. 9.

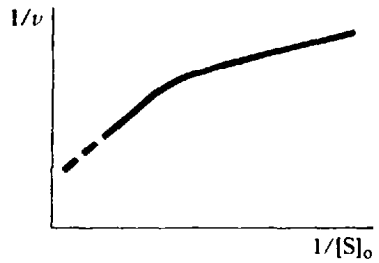


Fig. 6-28

SIGNALING

6.10. Extracellular epinephrine (adrenaline) (from the adrenal medulla) activates β_3-adrenergic receptors on fat cells to induce the breakdown of triacylglycerols to free fatty acids and glycerol. The intracellular enzyme involved in this process, hormone-sensitive lipase, is activated by protein kinase A. What are the key elements of the signal transduction cascade?

SOLUTION

Adrenergic receptors belong to the G-protein coupled receptor family. The key elements are: G_s, adenylate cyclase, 3',5'-cyclic AMP and protein kinase A.

6.11. Which receptors are fastest acting and which receptors act slowly?

SOLUTION

In general, the fastest-acting receptors are those that involve the least number of molecular steps. Ligand-gated ion channels respond to their activators within milliseconds. G-protein coupled receptor-dependent effects are generally observed within seconds. Nuclear receptors require gene transcription, the translation of mRNA into proteins, post-translational processing and modification and, in some cases, cell cycle turnover. As a result, the effects of these receptors are not evident for hours or even days. Growth factor receptors provoke early as well as late events. The early events arise as a result of the activity of cellular enzymes, e.g., some phospholipid modifying enzymes. The late events arise as a result of effects on gene transcription.

6.12. Human neutrophils were exposed to the tumor promoter phorbol myristate acetate (PMA, 100 n*M*). After a steady state of metabolism had been achieved, the cell suspension was homogenized and, after centrifugation, membrane and supernatant fractions were collected. Membrane and supernatant fractions were also prepared from cells that had not been exposed to PMA. Samples of both homogenates (i.e., prior to centrifugation) were also collected. The fractions were then assayed for the presence of protein kinase C. After the protein kinase C assay was performed the following data were obtained:

	Homogenate %	Membranes %	Supernatant %
Control cells	100	10	90
PMA-treated cells	100	60	40

Note: all data are expressed as percentages with respect to the control homogenate.

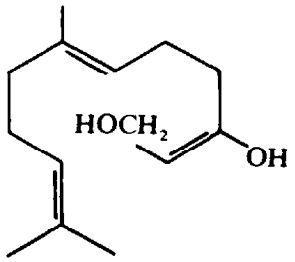
What is the simplest explanation of these data?

SOLUTION

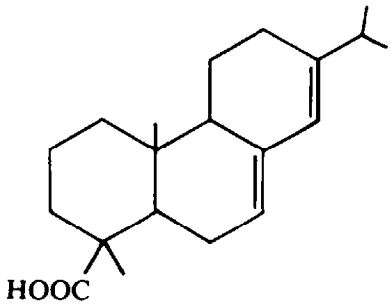
After PMA treatment the recovery of protein kinase C in the homogenate is 100% when compared to the control homogenate. However, a redistribution of protein kinase C has occurred. Upon activation of protein kinase C by PMA (or by its natural activators DAG and Ca^{2+} ions) protein kinase C is translocated to the plasma membrane where it phosphorylates protein targets.

Supplementary Problems

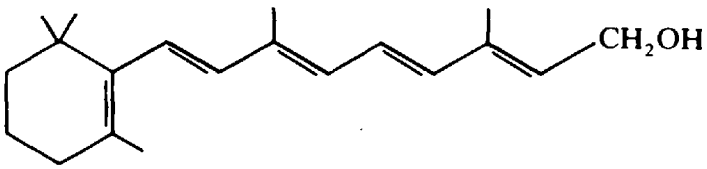
6.13. Indicate the ways in which isoprene units are linked together in forming the following compounds:



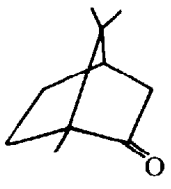
(*a*) Farnesol



(*b*) Abietic acid



(*c*) Vitamin A



(*d*) Camphor

6.14. (*a*) How would you define a *glycolipid*? (*b*) Give two examples.

6.15. Why do almost all naturally occurring fatty acids contain an even number of carbon atoms?

6.16. Write the structures for the following fatty acids:

(*a*) Myristic acid (14:0)

(*b*) Myristoleic acid ($14{:}1^{\Delta 9}$)

(*c*) Ricinoleic acid ($18{:}1^{\Delta 9}$ C-12 hydroxylated)

6.17. How many different molecules of triacylglycerol can be made from glycerol and four different fatty acids?

6.18. Write the structures for two triacylglycerols, one of which is (*a*) solid and the other (*b*) liquid at 37°C.

6.19. Electrophoresis of the following mixture of lipids was carried out at pH 7: PE, PS, PG, DPG. State whether these lipids would move toward the anode or the cathode or would remain stationary.

6.20. Draw a structure showing the conformation of cholesterol.

6.21. A solution of 1-palmitoyl-2-stearoyl-3-myristoylglycerol and phosphatidic acid in benzene is shaken with an equal volume of water. After the two phases separate, which lipid will be in the higher concentration in the aqueous phase?

6.22. How many phospholipid molecules are there in a 1-μm^2 region of a phospholipid bilayer? Assume that a phospholipid molecule occupies 0.7 nm^2 of the surface area.

6.23. The lipids isolated from the membranes of 4.74×10^9 human red blood cells were spread as a monolayer with an area of 0.89 m^2. Assuming the red blood cell approximates a disk 7 μm in diameter and 1 μm thick, show that the membrane covering the red blood cell must be two lipid molecules thick.

6.24. Why does phosphatidylethanolamine partition preferentially into the inner leaf of artificial vesicles composed of PE and PC?

6.25. Predict the effects of the following operations on the phase-transition temperature and on phospholipid mobility in vesicles made from dipalmitoylphosphatidylcholine:

(*a*) Introducing dipalmitoleoylphosphatidylcholine into the vesicles

(*b*) Introducing a high concentration of cholesterol into the vesicles

(*c*) Introducing integral membrane proteins into the vesicles

6.26. Calculate the average density of a membrane composed of 30 percent (by weight) protein (density, 1.33 $g\,cm^{-3}$) and 70 percent by weight phosphoglyceride (density, 0.92 $g\,cm^{-3}$).

6.27. For the membrane referred to in Prob. 6.26, how many molecules of lipid are there for each molecule of protein? Assume an average M_r of 800 for the phosphoglyceride and 40,000 for the protein.

6.28. What will be the order of simple diffusion of the following compounds through a biological membrane: propionic acid, 1,3-propanediol, propionamide, 1-propanol, alanine?

6.29. In the bacterium *E. coli*, glucose is taken up by group translocation, lactose is taken up by secondary active transport (using H^+), and maltose is taken up by means of a binding-protein system. Outline how it would be possible to determine whether melibiose (a disaccharide of glucose and galactose) is taken up by *E. coli* and, if it is, whether one of the mechanisms described earlier applies.

6.30. A number of identical cell suspensions are treated with different amounts of radioactively labeled leucine. The initial rates of leucine uptake are measured for each suspension (see the table below). What is the maximum possible rate of uptake of leucine by cells using the same transport system?

Leucine Concentration (μM)	Initial Rate of Uptake (cpm)*
0.5	55
1	110
5	480
10	830
20	1,300
30	1,700
50	2,100
100	2,600

*cpm denotes counts per minute, a measure of radioactivity.

6.31. The pH of gastric juice is 1. The cells that produce gastric juice have an internal pH of 7. Calculate the ΔG for transport of protons from these cells into the stomach at 37°C. (R, the gas constant, is $8.3\ \text{J mol}^{-1}\ \text{K}^{-1}$.)

6.32. Predict the effects of (*a*) valinomycin and (*b*) gramicidin A on the initial rate of glucose transport into vesicles derived from cells that accumulate glucose by cotransport with Na^+. Assume that the outside medium contains $0.2\ M\ Na^+$ and that the interior of the vesicle contains an equivalent concentration of K^+.

6.33. In platelets, protein kinase C activates the enzyme inositol 1,4,5-trisphosphate-5-phosphomonoesterase, which breaks down inositol 1,4,5-trisphosphate (IP_3) to inositol 1,4-bisphosphate (IP_2). What effect would this have on Ca^{2+} release due to receptors that stimulate phospholipase C?

Chapter 7

Nucleic Acids

7.1 INTRODUCTION

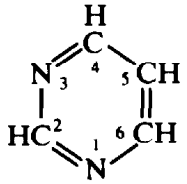
In 1868 Friederich Miescher isolated a substance from the nucleus of pus cells. It was considered to be characteristic of the nucleus, and he called it *nuclein*. A similar substance was subsequently isolated from salmon sperm heads. Nuclein was later shown to be a mixture of a basic protein and a phosphorus-containing organic acid, now called *nucleic acid*.

7.2 NUCLEIC ACIDS AND THEIR CHEMICAL CONSTITUENTS

The major nucleic acid in the nucleus of cells is *deoxyribonucleic acid* (or *DNA*). It contains the pentose sugar *deoxyribose* as one of its chemical constituents. DNA is now known to be the genetic material. Another type of nucleic acid, *ribonucleic acid* (or *RNA*), contains *ribose* instead of deoxyribose. Its main role is in the transmission of the genetic information from DNA into protein.

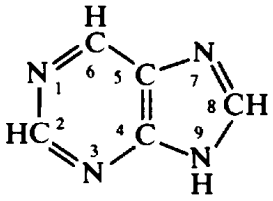
DNA molecules are very large, much larger than proteins. RNA is more comparable to proteins in size. Complete hydrolysis of DNA (or RNA) by acid cleaves it into a mixture of nitrogenous bases, 2-deoxy-D-ribose (or D-ribose for RNA), and orthophosphate. There are two general types of nitrogenous bases in both DNA and RNA, *pyrimidines* and *purines*.

Pyrimidines are derivatives of the heterocyclic compound *pyrimidine*:



Pyrimidine

Purines are derivatives of the fused-ring compound *purine*:



Purine

The numbering of the positions in the rings has been established by convention (IUPAC).

Pyrimidines

The major pyrimidines found in DNA are *thymine* and *cytosine*; in RNA, they are *uracil* and *cytosine*. These three pyrimidines differ in the types and positions of chemical groups attached to the ring.

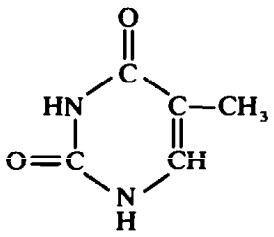
Thymine is 5-methyl-2,4-dioxypyrimidine.

Cytosine is 2-oxy-4-aminopyrimidine.

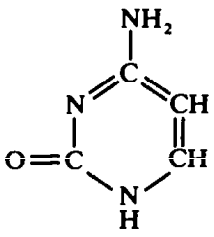
Uracil is 2,4-dioxypyrimidine.

EXAMPLE 7.1

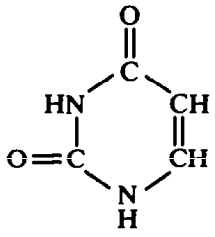
Write the structures of (*a*) thymine, (*b*) cytosine, and (*c*) uracil.



(*a*) Thymine



(*b*) Cytosine

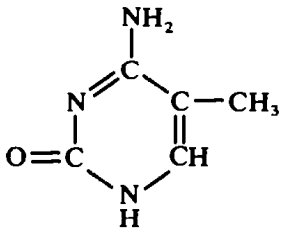


(*c*) Uracil

Thymine can also be described as 5-methyluracil. Other methylated pyrimidines are found in some nucleic acids.

EXAMPLE 7.2

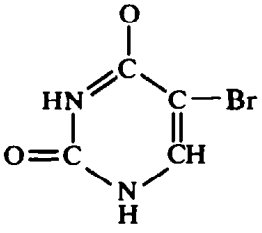
Write the structure of 5-methylcytosine.



5-Methylcytosine

Methylation of cytosine in both DNA and RNA has important biological implications with respect to protection of the genetic material and its expression.

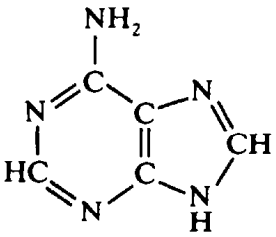
5-Bromouracil is an analog of thymine, differing only in the substituent on C-5 (Br instead of CH_3). These two substituents occupy approximately the same space, and the enzyme responsible for making DNA can accommodate either, allowing 5-bromouracil to be incorporated into DNA in certain types of cells and viruses. This has been of considerable value in studies of DNA synthesis.



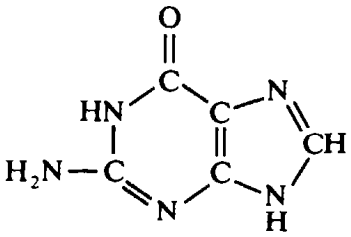
5-Bromouracil

Purines

The major purines found in DNA and RNA are *adenine* and *guanine*. They differ in the types and positions of chemical groups attached to the purine ring, as shown below:



Adenine



Guanine

EXAMPLE 7.3

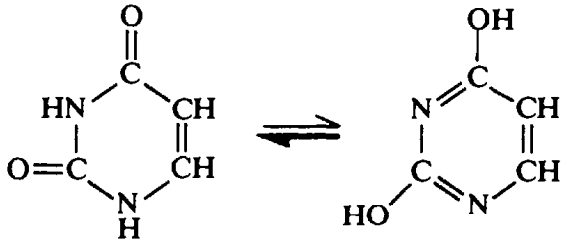
Describe adenine and guanine in terms that indicate the nature and positions of substituent chemical groups on the purine ring.

Adenine is 6-aminopurine.

Guanine is 6-oxy-2-aminopurine.

Tautomeric Forms of Pyrimidines and Purines

All pyrimidines and purines can exist in alternative isomeric forms called *tautomers*. Thus, uracil can exist in *keto* and *enol* forms.

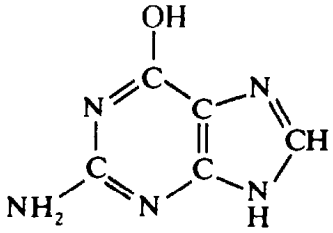


Keto Enol

The heavy arrow indicates that the keto form is strongly preferred at neutral pH.

EXAMPLE 7.4

Write the enol form of guanine.



Enol form of guanine

Question: Is it possible to write an enol form for adenine?

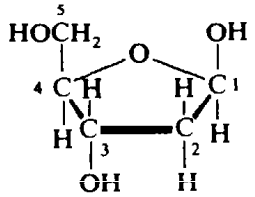
No, because it does not contain keto groups. It can, however, isomerize to the tautomeric *imino* form, but the *amino* form shown earlier in this section predominates.

Sugars

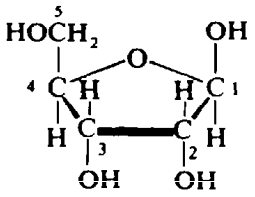
The sugar in DNA is 2-deoxy-D-ribose; in RNA it is D-ribose.

EXAMPLE 7.5

Write the forms of these sugars as they occur in DNA and RNA.



2-Deoxy-β-D-ribose
(2-deoxy-β-D-ribofuranose)



β-D-Ribose
(β-D-ribofuranose)

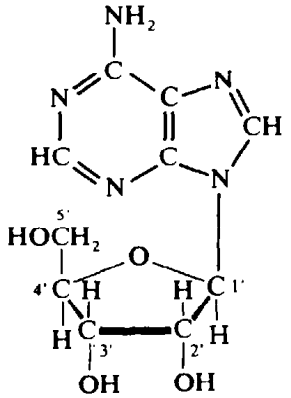
Note that it is the β anomer in each case that is present in the nucleic acid (Chap. 2).

7.3 NUCLEOSIDES

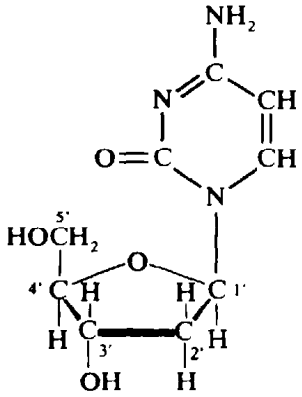
Within the structure of the nucleic acids, a pyrimidine or purine is linked to the sugar (2-deoxy-D-ribose or D-ribose) to give a *nucleoside*. The nucleosides are referred to as *deoxyribonucleosides* if they contain deoxyribose, and *ribonucleosides* if they contain ribose. The purine nucleosides have a *β-glycosidic linkage* from N-9 of the base to C-1 of the sugar. In pyrimidine nucleosides, the linkage is from N-1 of the base to C-1 of the sugar.

EXAMPLE 7.6

Write the structures of (*a*) the ribonucleoside containing adenine and (*b*) the deoxyribonucleoside containing cytosine.



(*a*) 9-β-D-Ribofuranosyl adenine (adenosine)



(*b*) 1,2′-Deoxy-β-D-ribofuranosyl cytosine (deoxycytidine)

Because the glycosidic linkage in Example 7.6 is to a nitrogen in the pyrimidine or purine, these nucleosides are referred to as N-*glycosides*. To distinguish the atoms in the furanose ring of the sugar from those in the rings of the bases, the former are designated 1′, 2′, . . ., 5′, as shown. The chemical

names written immediately below the structures are concise but awkward to use, and it is more convenient to use simpler terms. Thus:

Adenine linked to ribose ≡ adenosine

Uracil linked to ribose ≡ uridine

Guanine linked to ribose ≡ guanosine

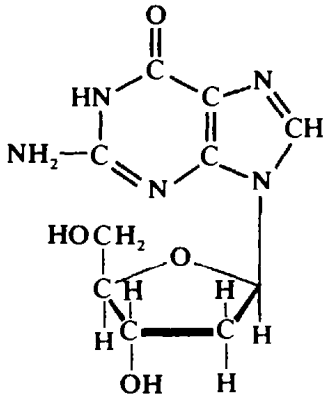
Guanine linked to deoxyribose ≡ deoxyguanosine

Cytosine linked to deoxyribose ≡ deoxycytidine

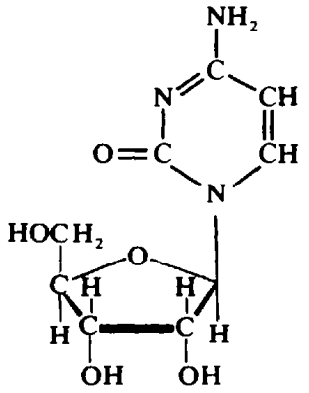
Thymine linked to deoxyribose ≡ deoxythymidine

EXAMPLE 7.7

Write the structures of (*a*) deoxyguanosine and (*b*) cytidine.



(*a*) Deoxyguanosine



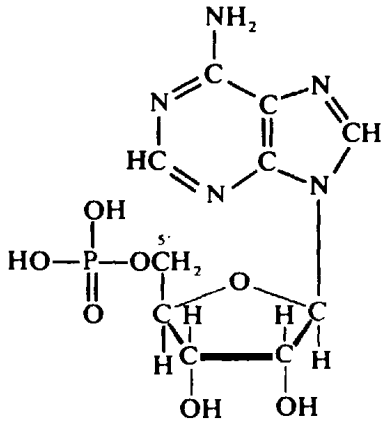
(*b*) Cytidine

7.4 NUCLEOTIDES

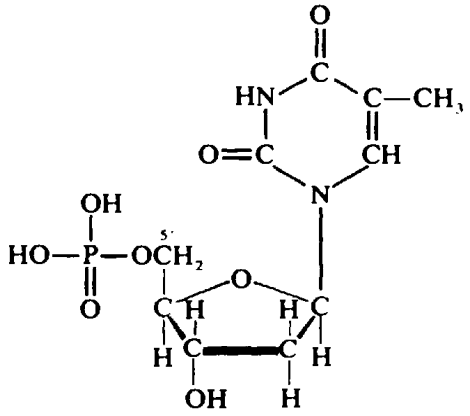
The *nucleotides* are phosphoric acid esters of nucleosides, with phosphate at position C-5′. Nucleotides with phosphorylation at other positions are known, but they are not components of the nucleic acids. Nucleotides containing deoxyribose are called *deoxyribonucleotides*; those containing ribose are known as *ribonucleotides*.

EXAMPLE 7.8

Write the structures of (*a*) the ribonucleotide containing adenine and (*b*) the deoxyribonucleotide containing thymine.



(*a*) Adenosine 5′-phosphate (AMP)



(*b*) Deoxythymidine 5′-phosphate (dTMP)

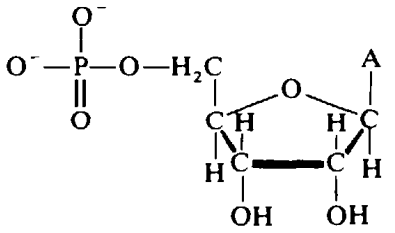
Adenosine 5′-phosphate is also known as *AMP* (for adenosine monophosphate) or *adenylic acid*. If deoxyribose replaces ribose in adenosine 5′-phosphate, the terminology is *dAMP* or *deoxyadenylic acid*. The abbreviated names for some ribonucleotides and deoxyribonucleotides are listed below.

Base	Ribonucleotide	Deoxyribonucleotide
Adenine, A	Adenylic acid, AMP	Deoxyadenylic acid, dAMP
Guanine, G	Guanylic acid, GMP	Deoxyguanylic acid, dGMP
Cytosine, C	Cytidylic acid, CMP	Deoxycytidylic acid, dCMP
Uracil, U	Uridylic acid, UMP	Deoxyuridylic acid, dUMP
Thymine, T	Thymidylic acid, TMP	Deoxythymidylic acid, dTMP

The terminology tells us that the nucleotides are acids. This results from the primary phosphate ionization, which has a pK_a value of approximately 1 (Chap. 10). The nucleotides are thus negatively charged at neutral pH; also contributing to this negative charge is the ionization of the secondary phosphate, which has a pK_a value of approximately 6. At neutral pH, there is no charge on any of the bases.

EXAMPLE 7.9

Write the structure of the charged form of AMP at pH 7.



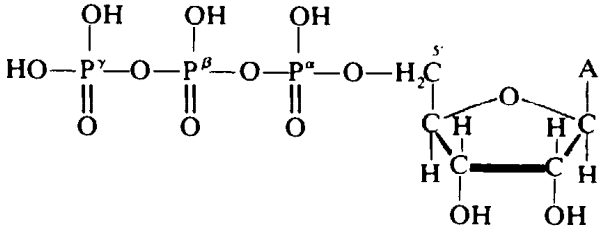
AMP (charged form)

Note that two negative charges reside on the phosphate. In this structure, adenine is represented by A; it carries no charge.

All the common 5′-nucleotides exist also as 5′-diphosphates and 5′-triphosphates. These contain two and three phosphates, respectively. The corresponding adenosine 5′-nucleotides are referred to as *ADP* and *ATP*.

EXAMPLE 7.10

Write the structure of ATP.



ATP

Note that this has been represented as an uncharged form. The phosphorus atoms are designated α, β, and γ, the α phosphorus being attached to the 5′ C of the ribose. *ADP* contains only α and β phosphates.

Question: What is the net charge on ATP at neutral pH?

It is negative and in the range −2 to −4. This is because ATP, and all nucleoside triphosphates, can dissociate four protons from the phosphate groups. The first has a pK_a of ~1, and the second, third, and fourth in the range 6 to 7, approximately.

EXAMPLE 7.11

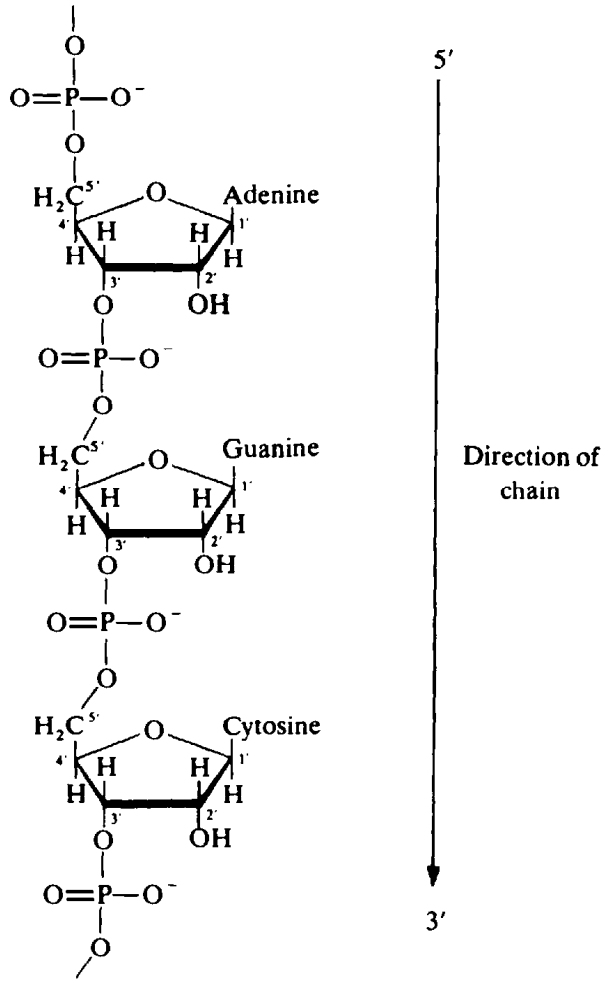
Arrange ATP, dAMP, and CDP (cytidine *di*phosphate) in order of increasing net negative charge at pH 7.

The order is dAMP, CDP, ATP, because the phosphates have the dominating effect at neutral pH, and the more phosphates there are in a molecule, the greater will be its negative charge.

The ribonucleoside di- and triphosphates (NDPs, NTPs) and deoxyribonucleoside di- and triphosphates (dNDPs, dNTPs) have important functions in the cell. They operate as *energy carriers* in various reactions and as precursors for the synthesis of nucleic acids (Chaps. 10 and 16).

7.5 POLYNUCLEOTIDES

The nucleic acids, both DNA and RNA, are *polynucleotides*; that is, they are polymers containing nucleotides (various types) as the repeating subunits. The nucleotides are joined to one another through *phosphodiester linkages* between the 3′ C of one nucleotide and the 5′ C of the adjacent one. This linkage is repeated many times to build up large structures (chains or strands) containing hundreds to millions of nucleotides within a single giant molecule.

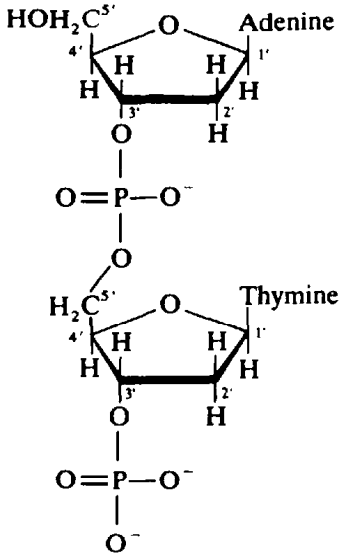


EXAMPLE 7.12

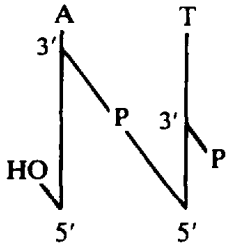
Write the structure (below) of a section of a poly*ribo*nucleotide (RNA) chain containing adenine, guanine, and cytosine as a sequence of nucleotides. Note that the structure is written in its charged form. Because the phosphodiester linkages join different carbons, 3′ and 5′, in adjacent nucleotides, the chain has a chemical direction, or *polarity*. By convention, the structure shown here has the direction 5′ → 3′ downward (or 3′ → 5′ upward). Also, the order (or sequence) of nucleotides is conventionally written in the 5′ → 3′ direction. Thus, the section shown has the sequence adenine, guanine, cytosine; or AGC, for short. To indicate that there are phosphates attached at the 5′ and 3′ ends of the structure shown, it is more accurately referred to as pApGpCp. This shorthand form does not indicate that the sequence of three nucleotides is just a portion of a much longer structure; pApGpCp could also refer to a molecule containing only three nucleotide units (a trinucleotide) in which the 5′ and 3′ ends are phosphorylated. If the sequence AGC had been present in a poly*deoxy*nucleotide, the structure would be written in the shorthand form as d-pApGpCp, or just dAGC.

EXAMPLE 7.13

Write the structure of the dinucleotide d-ApTp.



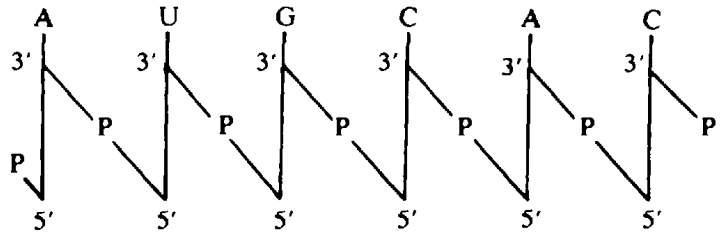
Note the absence of a phosphate at the 5′ end and the presence of deoxyribose. Another commonly used shorthand form for describing this structure is



Note that this latter form does not define the nature of the sugar.

EXAMPLE 7.14

Write the structure of pApUpGpCpApCp in the shorthand form.



This structure contains six nucleotide units and is referred to as a *hexanucleotide*. The general term for structures containing a few nucleotides (10 or less) is *oligonucleotide*.

7.6 STRUCTURE OF DNA

DNA is a *polydeoxynucleotide* and among the largest of the biological macromolecules; some DNA molecules comprise more than 10^8 nucleotides. They contain adenine, thymine, guanine, and cytosine as the bases, and the genetic information is encoded within the nucleotide sequence, which is precisely defined over the entire length of the molecule. One of the simplest methods for determining the nucleotide sequence of DNA makes use of an enzyme, *DNA polymerase*, which catalyzes the synthesis of DNA. The properties of this enzyme are discussed in Chap. 16.

Base Composition of DNA

The base composition of DNA from many different species has been determined. It varies from one to another (see Table 7.1).

Question: What is the base composition of DNA from human kidney?

It is the same as for human liver, as shown in Table 7.1, because the base composition of DNA is a characteristic of a particular species and does not vary from one cell type to another. This reflects the fact that the nucleotide sequence, and therefore the genetic information present, in each type of cell within an organism is exactly the same. However, as will be seen later, this information is expressed differently in the various cell types of an organism (Chap. 17).

Table 7.1. Base Composition of DNA in Various Species

	Base Composition (mol %)			
Species	G	A	C	T
Sarcina lutea	37.1	13.4	37.1	12.4
Alcaligenes faecalis	33.9	16.5	32.8	16.8
E. coli K12	24.9	26.0	25.2	23.9
Wheat germ	22.7	27.3	22.8*	27.1
Bovine thymus	21.5	28.2	22.5*	27.8
Human liver	19.5	30.3	19.9	30.3
Saccharomyces cerevisiae	18.3	31.7	17.4	32.6
Clostridium perfringens	14.0	36.9	12.8	36.3

*Cytosine + methylcytosine.

EXAMPLE 7.15

Are there any features common to DNA from various species with respect to the ratio of one base (or type of base) to another?

The ratio of purines (A + G) to pyrimidines (T + C) is close to unity in all cases. Perhaps more remarkable is that the ratios of both A to T and G to C are each close to unity. These two facts reflect an important structural feature of most DNAs.

Double-Helical Structure of DNA

Question: What structural feature of DNA accounts for the ratio of A to T and G to C being close to unity?

DNA is a duplex molecule in which two polynucleotide chains (or strands) are linked to one another through specific *base pairing* (Fig. 7-1). Adenine in one strand is paired to thymine in the other, and guanine is paired to cytosine. The two chains are said to be *complementary*. This was one of the essential features of Watson and Crick's proposal regarding the structure of DNA. Hydrogen bonds form between the opposing bases within a pair. In the structure proposed by Watson and Crick, A : T and G : C base pairs are roughly planar, with H bonds (dotted lines), as shown in Fig. 7-1. Note that two H bonds form in an A : T pair and three in a G : C pair.

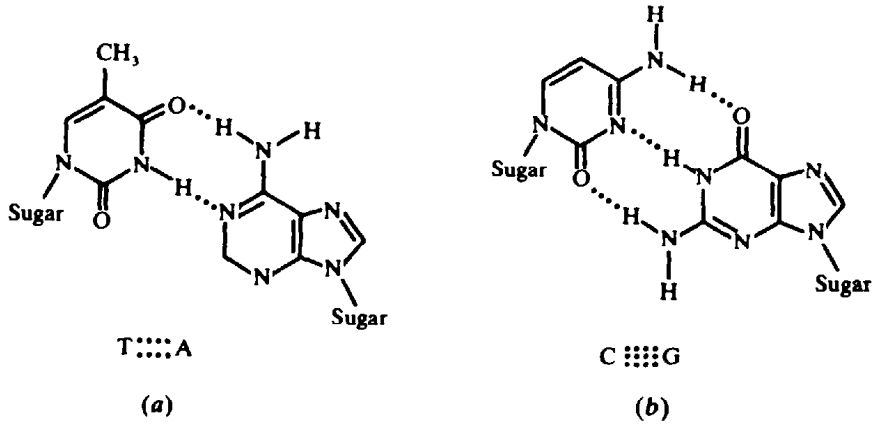


Fig. 7-1 Base pairing in DNA.

The base pairs are stacked on top of one another, with the plane of the base pairs being perpendicular to the length of the duplex. This is shown diagramatically in the ladder-type structure in Fig. 7-2.

Question: In the duplex, ladder-type structure shown in Fig. 7-2, why are the two chains orientated in opposite directions?

A model for DNA incorporating base pairing between complementary strands and consistent with x-ray diffraction data was developed by J. Watson and F. Crick in 1953. Basic to the structure was the twisting of the two strands around one another to give a *right-handed* helix (the *double helix*), and to achieve a structure consistent with data available at the time, it was necessary to orient the complementary chains in opposite directions (Fig. 7-3). Direct proof for this *opposite polarity* in chain direction was achieved about 10 years later.

Question: How does the twisting into a helix contribute to the stability of the overall structure of DNA?

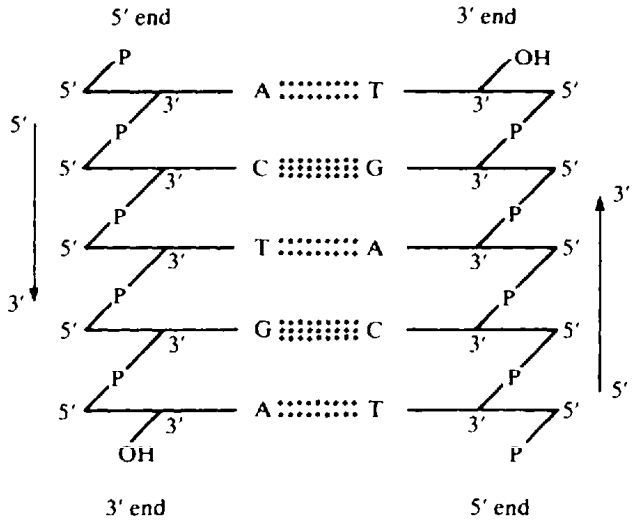


Fig. 7-2 Stacking of base pairs within the DNA duplex. Each group of dotted lines, representing the H bonds between the base pairs, is in a plane perpendicular to the surface of the page.

One of the most significant effects of twisting into a helix is to bring the stacked base pairs very close to one another. In the B form (see below) of the helix this distance, which is called the *rise*, is 0.34 nm. Consequently, water is excluded from what is now a hydrophobic core; the charged phosphates are on the surface. The hydrophobic interactions within the core contribute, along with the H bonds between the base pairs, to the overall stability of the helix. It should also be noted that there is one complete twist of the helix every 10 base pairs or 3.4 nm. This distance is referred to as the *pitch* of the helix. The surface of the helix shows alternating *major* and *minor grooves*, which follow the twist of the double-stranded molecule along its length. The major groove is now known to accommodate interactions with proteins that recognize and bind to specific nucleotide sequences.

Question: The structure of the double helix shown in Fig. 7-3 is called the *B form*. What are the other forms, and do they have a biological role?

The x-ray diffraction data used by Watson and Crick was obtained from fibers of DNA that were prepared under conditions of high humidity. At lower humidities (<75 percent), a fiber of DNA will shorten. This is a result of a change to the *A form*, in which the base pairs are not perpendicular to the helix axis; they are tilted about 20°, and the pitch is reduced to 2.8 nm with 11 base pairs per turn. While the B-form of the helix is the predominant form in cells, double-stranded RNA and DNA/RNA hybrid molecules generate the A form of the helix.

A dramatically different form of the double helix has been observed in DNA containing alternating purine–pyrimidine sequences, especially d(CG)*n* but also d(TG)*n*. It is a left-handed, rather than a right-handed helix and is known as the Z form of DNA. Helix parameters are contrasted in Table 7.2 and space-filling models of the Z form and B form of DNA are compared below (Fig. 7-4.).

In the Z form, the repeating unit is a *dinucleotide* and the resultant structure has the staggered zig-zag shape of the sugar-phosphate backbone from which the name (Z) was derived. It is possible that the Z form of DNA does have an important biological role although at present this is uncertain. Though sequences of $d(TG)_n$ where $n > 25$ are common in eukaryotic DNA ($\approx 10^5$ copies in the human genome), it has not been shown if they adopt the Z form *in vivo*.

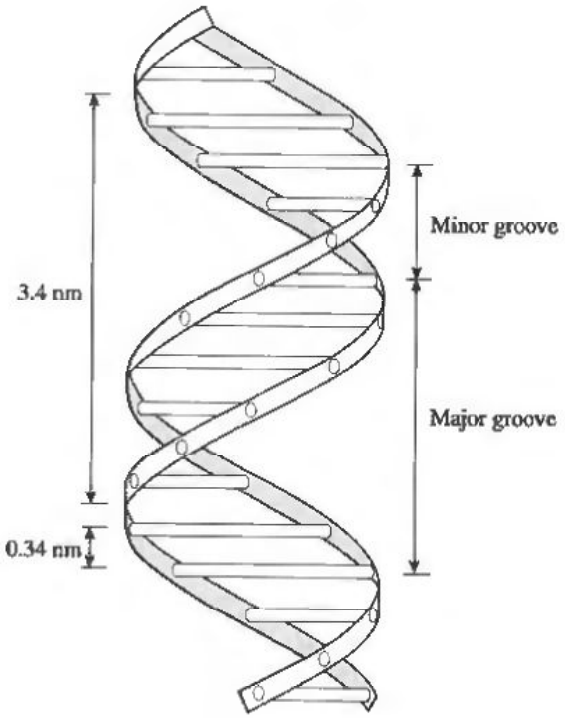


Fig. 7-3 Diagram of the DNA double helix (the so-called B form).

Table 7.2. Comparison of DNA Helices

Helix Attribute	A-DNA	B-DNA	Z-DNA
Sense	right	right	left
Base pairs per turn	11	10	12
Rise (nm)	0.29	0.34	0.37
Pitch (nm)	3.2	3.4	3.5

Local Variation in Helix Structure

The x-ray diffraction of DNA fibers can only produce data which averages the effect of sequence variation on DNA conformation. Higher resolution studies require the x-ray diffraction analysis of single DNA crystals and this became possible only with the chemical synthesis of pure DNA of defined length and sequence.

The examination of crystallized DNA fragments of varying base composition show *sequence-dependent* variations in the double helix while emphasizing the A, B and Z forms as structurally distinct families. Variation occurs in the orientation of each base pair to the next by rotation about the *X* axis or *tilt*, the *Y* axis or *roll* or the *Z* axis or *twist* of the helix. Bases in a pair may also rotate in opposition producing variations in *buckle*, *propeller* and *opening* (Fig. 7-5).

The parameters listed in Table 7.2 should therefore be considered average values; for example, although the mean rise for the B form has been measured as 0.34 nm this parameter has been shown to vary from 0.25 nm to 0.44 nm.

Question: What sequence-dependent variations have been described? Stretches of dA_n, where $n = 4$–6, occurring in phase with the helical repeat (every ≈10–11 base pairs), lead to curvature of

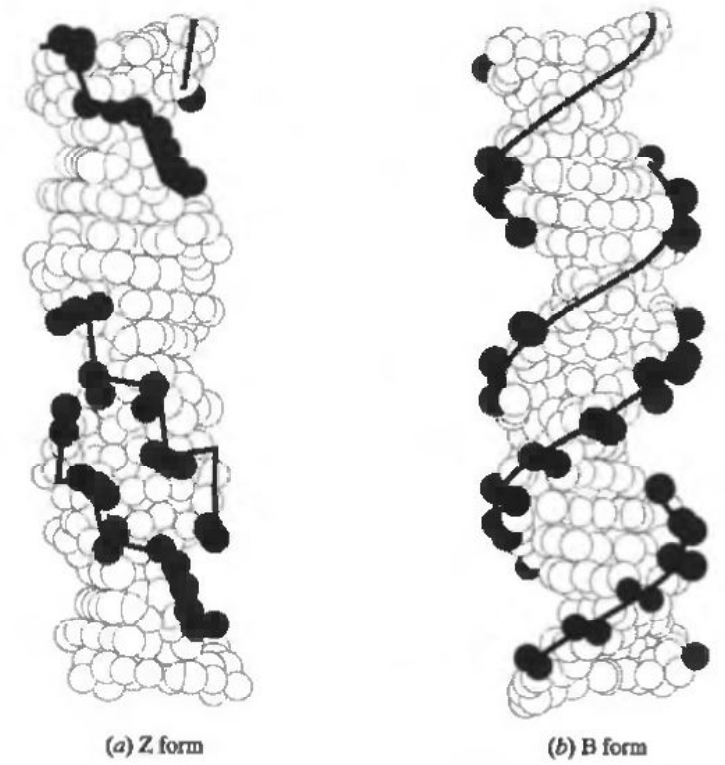


Fig. 7-4 The Z form and B form of DNA.

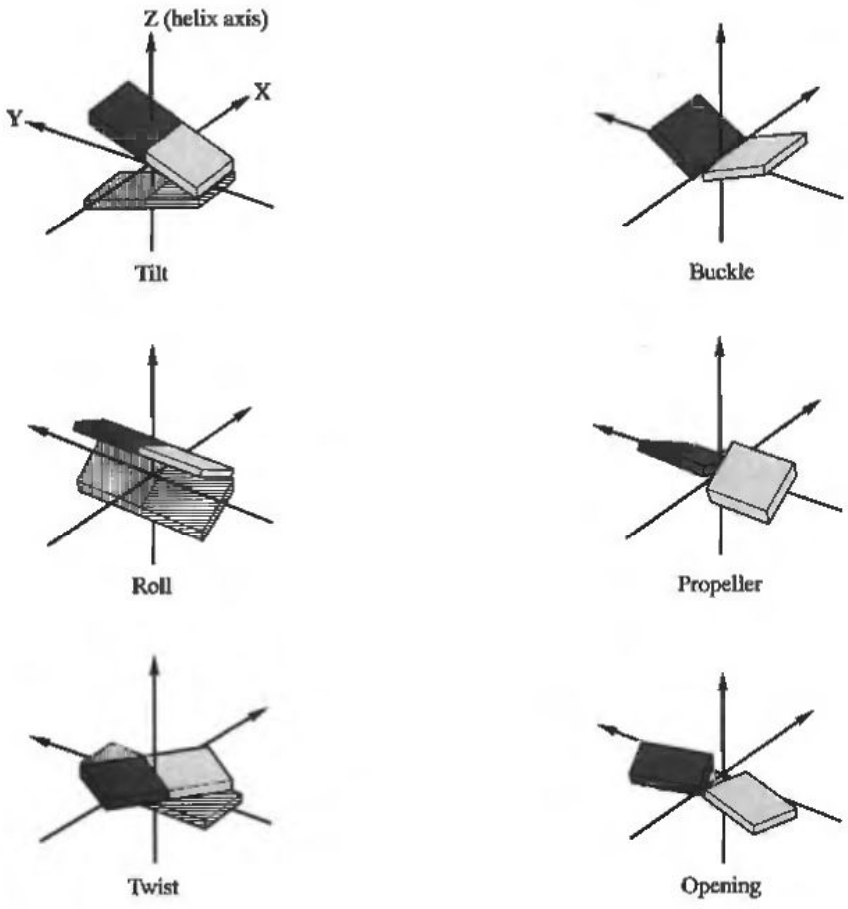


Fig. 7-5 Sequence-dependent variations in the base-pairing geometry of DNA.

the helical axis of the B form. Similarly purine-pyrimidine-x-x-x-pyrimidine-purine-x-x-x repeating sequences, where x is any base, lead to curvature.

The discovery of the Z form of the helix resulted from the study of crystals formed from the sequence dCGCGCG (see above). Sections of DNA exhibiting this alternating purine-pyrimidine sequence may lead to a "wrinkling" of the B form of the helix, producing a deeper minor groove. Alternatively there may be a transformation to the Z form of the helix.

Question: How are oligonucleotides synthesized in the laboratory?

A common method for DNA synthesis is the phosphite triester method (Fig. 7-6). A single-stranded oligonucleotide is formed by the sequential creation of diester bonds between the 5′-hydroxyl of one residue and the 3′-phosphate of the next. The 3′-phosphate is activated by substitution by dialkyl phosphoamidite (DPA) and reacts readily with the free 5′-OH of the first nucleotide. To prevent the formation of unwanted linkages the first nucleotide is linked, by the 3′-OH, to a solid support (often silica gel) in a column or funnel; the 5′-OH of the second nucleotide is prevented from reacting by a dimethoxytrityl group (DMT).

The reaction results in a phosphite triester which is oxidized to a phosphodiester by flushing the column with iodine. Subsequently the DMT group is removed from the second nucleotide with 80% acetic acid (detritylation) in preparation for reaction with the next DPA-activated monomer. This cycle continues until the oligonucleotide is complete, and it is capable of generating sequences of up to 150 residues.

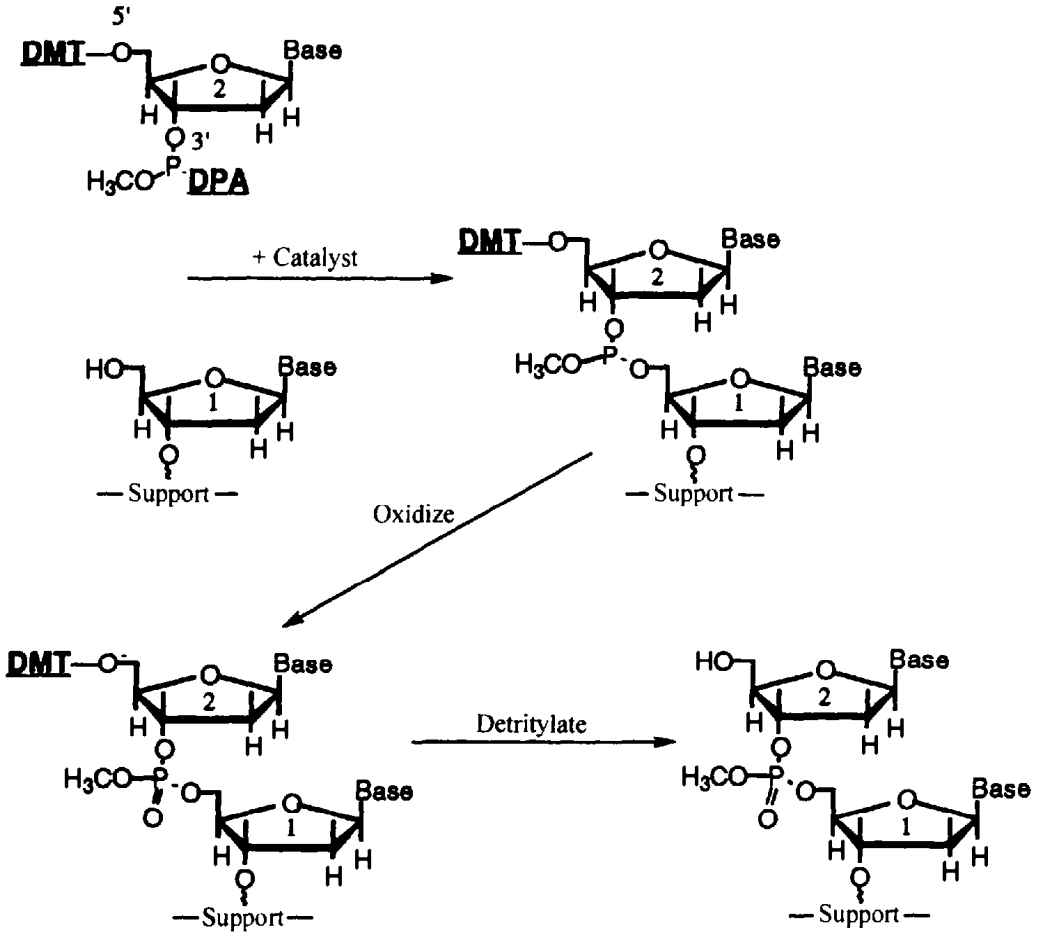


Fig. 7-6 Solid-phase oligonucleotide synthesis.

Question: How is DNA sequenced?

The most common method is called chain-termination sequencing or the dideoxy method, and is described in Chap. 16.

7.7 DENATURATION OF DNA

The double helix is a relatively stiff and elongated molecule. Consequently, a solution of DNA has a high viscosity. If such a solution is heated to ~95°C, the viscosity drops markedly, reflecting a collapse of the double-helical structure. This is known as *denaturation* and is accompanied by separation of the duplex into its single strands, which are fairly flexible. Denaturation and renaturation provide valuable information on important properties of the DNA obtained from various sources. Denaturation also provides the basis for very precise and sensitive approaches to the identification of specific sequences in both DNA and RNA. This has been central to the rapid developments in molecular genetics.

While denaturation can be detected readily through changes in viscosity, a much more convenient way to detect it is by ultraviolet (uv) absorption measurement. The difference in the uv absorption spectra of the native (double-helical) and denatured (single-stranded) forms of DNA is shown in Fig. 7-7. At the wavelength of maximum absorption (260 nm), absorption by single-stranded DNA is approximately 40 percent higher than by double-stranded DNA. This is referred to as the *hyperchromic effect* and results from the *unstacking* of the base pairs in the helix.

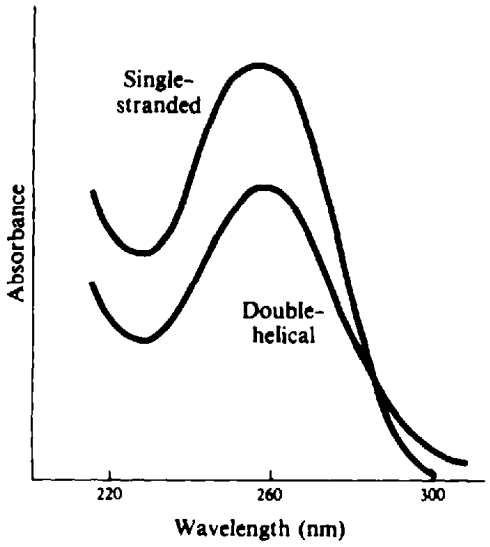


Fig. 7-7

Question: Will other treatments, in addition to heat, cause denaturation of DNA?

The DNA helix is stabilized by H bonds between individual base pairs as well as by hydrophobic forces between stacked base pairs. Reagents that reduce the H bonding and decrease the polarity of the surrounding medium, such as *formamide*, will cause denaturation. *Extremes of pH*, which endow the bases with a charge, are also effective. Thus, DNA at pH 12 shows absorption at 260 nm that is 40 percent higher than that of the native form.

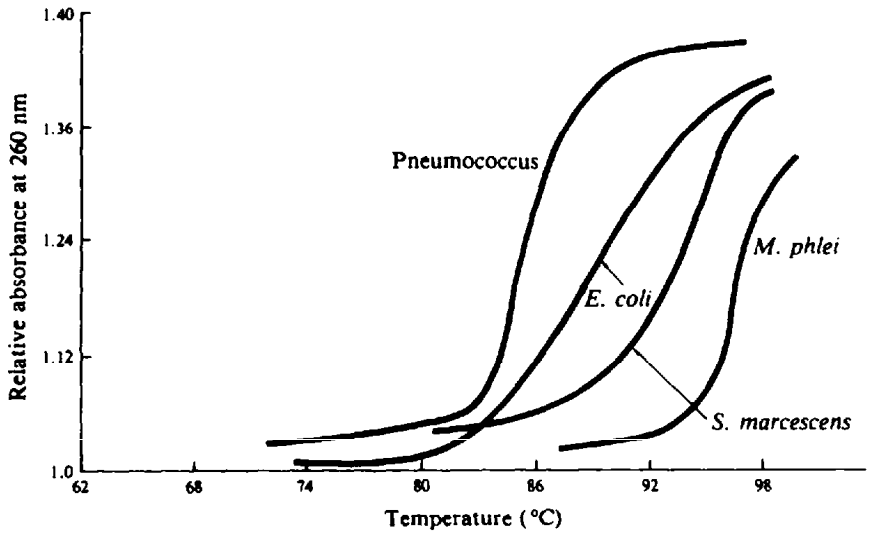


Fig. 7-8 Melting curves for DNA from different species.

Heat Denaturation of DNA

If the temperature of a solution of DNA is increased gradually, the change to the denatured form can be monitored by the change in absorbance at 260 nm. Typical results for several types of DNA are shown in Fig. 7-8.

The curves are referred to as *melting curves*, because the region over which the absorbance increases reflects the collapse (or *melting*) of the highly organized, semicrystalline state of double-helical DNA. The temperature at which 50 percent melting has occurred is called the *melting temperature*, or T_m.

EXAMPLE 7.16

What other factor besides heat affects the T_m of a particular DNA?

The T_m at neutral pH is dependent on the salt concentration (or *ionic strength*; Chap. 10) of the medium. The curves shown in Fig. 7-8 were obtained at an ionic strength of just above 0.15. If this were reduced by 90 percent, all T_m values would be lowered by about 20°C. This results from the additional negative charge and consequent greater electrostatic repulsion (which aids in disruption of the helix) within the DNA structure at the lower ionic strength.

Question: Why do the DNAs from various sources have different T_m values?

This is because the DNAs have different amounts of G : C and A : T base pairs, and the former confer the greater stability to the helix, perhaps through the presence of three H bonds per base pair rather than two (Fig. 7-1). Thus, the higher the GC content, the higher is the T_m. The value of T_m under standard conditions can be used, therefore, to obtain an estimate of the G + C content of an unknown DNA. This is obvious from Fig. 7-9, which shows a plot of T_m versus G + C content for a number of DNAs.

Renaturation of DNA

The complementary strands of DNA, separated by heat, spontaneously reassociate when the temperature is lowered below the T_m. This renaturation is also referred to as *annealing*.

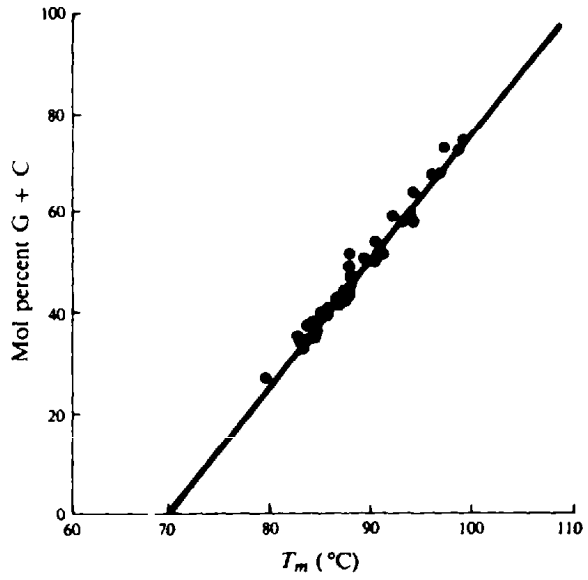


Fig. 7-9 Melting temperature of DNA as a function of the G + C content.

Question: Is the rate of annealing the same for all types of DNA?

The rate of renaturation depends on the concentration of complementary sequences. Viral DNA has a smaller variety of sequences than does bacterial DNA; this reflects the higher level of genetic complexity in bacteria. Thus, for viral and bacterial DNA fragments of the same average size and at the same molar concentration, there would be a higher concentration of complementary sequences in the former. Viral DNA therefore would renature faster than bacterial DNA. In other words, bacterial DNA has greater *sequence heterogeneity*.

Rates of renaturation and sequence heterogeneity (or complexity) can be examined quantitatively through *COT analysis*. If C_0 is the initial concentration of DNA (moles per liter DNA phosphate) and k is the rate constant for association of the complementary strands, it can be shown that the fraction f of single-stranded molecules decreases with time t (s) according to the expression

$$f = \frac{1}{1 + kC_0t} \tag{7.1}$$

It is usual to plot the results of a COT analysis as f versus C_0t. The behavior of several DNAs for a fixed set of conditions (size of DNA fragments, temperature, pH, ionic strength) is shown in Fig. 7-10. The value of C_0t when $f = 0.5$ is known as $C_0t_{1/2}$.

Question: What is the significance of the different values of $C_0t_{1/2}$?

The rate constant k is characteristic of a particular DNA and is related to its complexity in terms of sequence composition. $C_0t_{1/2}$ is the reciprocal of k and can therefore be used as a measure of sequence complexity. The higher the value of $C_0t_{1/2}$, the more complex is the DNA. Thus from the analysis shown above, bacterial DNA is more complex than viral DNA.

Question: What is meant by highly repeated and nonrepeated DNAs, shown in the above analysis?

The mouse highly repeated DNA is seen to have a $C_0t_{1/2}$ of $\sim 10^{-3}$ *M* s. It represents the most rapidly renaturing DNA of those examined. The mouse genome (see Sec. 7.8) contains about 10^6

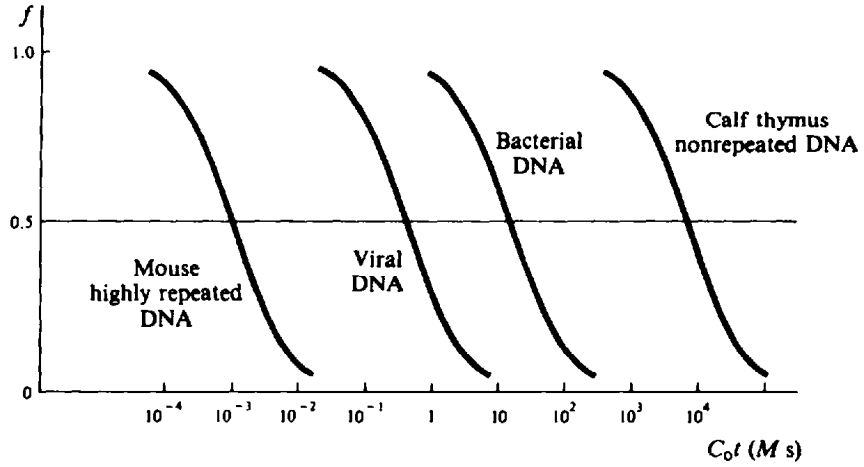


Fig. 7-10 COT analysis of various DNAs.

copies of a repeating sequence of ~300 base pairs; this is known as *highly repeated DNA*. Thus, this fraction of DNA is simple in structure, relatively high in molar concentration, and able to renature rapidly. At the other extreme, the calf thymus nonrepeated DNA has a very high $C_0t_{1/2}$. This reflects the reassociation of unique copies of sequences in a fairly complex genome. When total DNA from animal cells is examined in a COT analysis, it is usual to find a three-stepped curve resulting from highly repeated, moderately repeated, and unique (nonrepeated) sequences. The unique sequences are those that code for protein products. The highly repeated sequences are located in the centromeric region of chromosomes and could be involved in chromosome-chromosome recognition. Little is known about the moderately repeated sequences. Viral and bacterial DNA do not show multiple steps in a COT analysis and contain no highly or moderately repeated sequences.

7.8 SIZE, ORGANIZATION, AND TOPOLOGY OF DNA

DNA molecules are very long. For example the DNA in a bacterial cell is contained within a single double-helical molecule, which, when spread out, is about 1,000 times as long as the diameter of the rod-shaped cell. This molecule carries all the genetic information of the cell and thus describes the *genome* (a single complement of the genetic material).

The term *chromosome* refers to a physical or organizational unit within which part of or all the genome is contained. Thus, the *E. coli* genome is contained within just one chromosome, comprising a single DNA molecule. It has a size of 2.5×10^9 Da and contains approximately 4.6×10^6 base pairs. The size of DNA molecules is more commonly expressed in *kilobase pairs* (kb, 1,000 base pairs). The *E. coli* chromosome is 4,639 kb in size. Another feature of this particular molecule is that it is a closed, or "circular," structure, i.e., there are no free ends.

EXAMPLE 7.17

What variation in genome size, chromosome number, and DNA topology occurs among various organisms?

This information is summarized below for a number of commonly investigated viruses, bacteria, and eukaryotes (organisms whose cells contain nuclei).

Note that in Example 7.17, the DNA of bacteriophage ϕX174 is *single-stranded*, not double-stranded. In this case, the genome size given in kb refers to the number of base pairs in an equivalent duplex form. In progressing from the simple viruses to eukaryotes, the amount of information in the

Organism	Genome Size (kilobases)	Chromosomes per Genome	DNA Topology
Viruses			
Simian virus 40 (SV40)	5.1	1	Circular
Bacteriophage ϕX174	5.4	1	Circular, single-stranded
Bacteriophage λ	48.6	1	Linear
Bacteria			
Escherichia coli	4,639	1	Circular
Eukaryotes			
Yeast	13,500	17	Linear
Human	2,900,000	23	Linear

genome increases through many orders of magnitude. As a rough approximation, the number of kilobases in the genome of viruses and bacteria can be considered equivalent to the number of genes, each coding for a protein product. This equivalency of base number to genes does not exist in eukaryotes because of the presence of unexpressed sequences of bases (see Chap. 17). In viruses and bacteria (prokaryotes), circular DNAs are common. In eukaryotes, it is generally considered that each chromosome contains a single linear molecule of double-helical DNA.

Question: How are the very long DNA molecules condensed into more compact structures within the chromosomes?

In eukaryotes, DNA does not exist free. It is complexed with an approximately equal mass of basic proteins called *histones*. For a long time it was thought that bacterial DNA did not form such complexes. While histones are absent from bacteria, there is increasing evidence for the presence of histonelike proteins in them that enable condensation of their DNA into its compact *nucleoid* form.

There are five types of histones, designated H1, H2A, H2B, H3, and H4. They are of fairly low molecular weight (M_r = 11,000–23,000) and contain a large portion of the basic amino acids arginine and lysine. In each case there is an unusual distribution of amino acids along the single polypeptide chain; the basic amino acids tend to be clustered in one half, with the other half being relatively hydrophobic. In addition, the histones contain many modified amino acid side chains, e.g., methylated arginines and acetylated lysines.

The nucleoprotein complex that is formed is called *chromatin*. Chromatin can be isolated as fibers from nuclei. When the fibers are spread and examined under the electron microscope, they appear as "beads on a string." Most of the histone is contained in the beads, called *nucleosomes*. The nucleosome bead consists of a set of eight histones, two each of H2A, H2B, H3, and H4, around which approximately 200 base pairs of DNA are wrapped. Digestion of chromatin with *nucleases* (see below) yields a core particle still containing the eight histones but only 140 base pairs of DNA. The rest of the DNA, which is accessible to digestion, functions as a linker between the cores, and it is likely that histone H1 is associated with this linker. The structure of the core particles has been examined by x-ray diffraction. Figure 7-11 shows the arrangement of the DNA with the histone octamer. To compact the DNA further within the nucleus, the nucleosomes are organized into condensed higher-order structures.

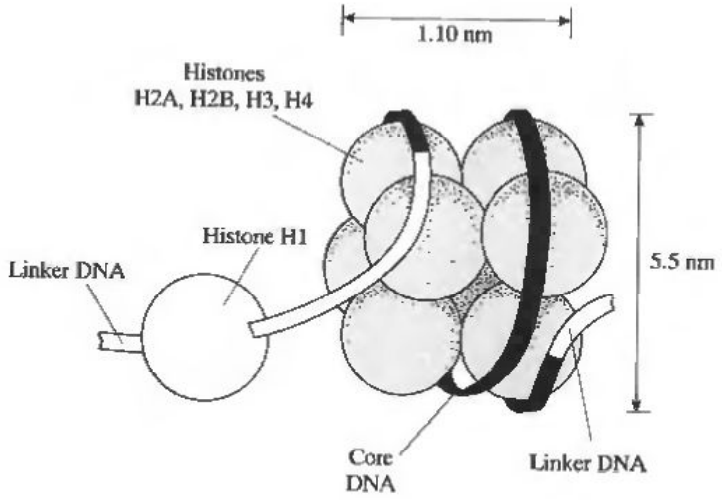


Fig. 7-11 The arrangement of histone and DNA in a nucleosome.

Question: What are plasmids?

Bacterial cells frequently contain additional DNA molecules called *plasmids*. These are relatively small (up to 200 kb) and are present in the form of circular duplexes. They can replicate independently of the bacterial chromosome and multiple copies can be present in a cell. Normally, they exist within a cell in a negatively supercoiled conformation, as does all DNA. Eukaryotic cells can also contain additional DNA to that present in the nucleus. Such *extrachromosomal* DNA is present in mitochondria and chloroplasts.

Question: What is meant by supercoiled DNA?

Supercoiling represents a twisting of the DNA double helix upon itself. It occurs in circular DNAs and in DNAs that are *topologically constrained* by being complexed to proteins. If a circular DNA molecule in a *relaxed* conformation (no supercoiling) is broken across both strands and one or more additional right-handed helical turns are inserted before rejoining the ends, the molecule would twist on itself to form a *positive supercoil*. If, on the other hand, the helix is unwound (given left-handed turns) before rejoining, the result is a *negative supercoil*. These forms, *topoisomers*, are shown in Fig. 7-12. Each supercoil depicted here contains a single *supertwist*. The number of supertwists in a molecule can be very large. Supercoiled DNA is readily converted into its relaxed form by the introduction of a break (or *nick*) between adjacent nucleotides in one of the two strands; this form is no longer topologically constrained.

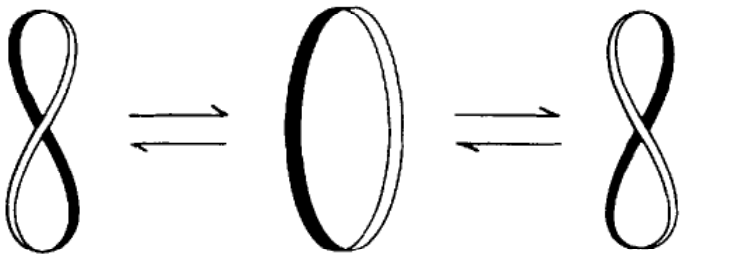


Fig. 7-12

7.9 STRUCTURE AND TYPES OF RNA

RNA comprises *polyribonucleotide* chains in which the bases are usually adenine, guanine, uracil, and cytosine. It is found in both the nucleus and the cytoplasm of cells.

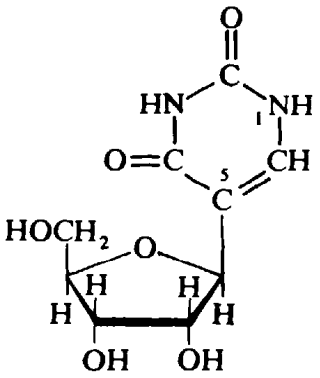
Question: What are the main differences, other than chemical composition, between RNA and DNA?

There are a greater variety of RNA forms, with molecular weights in the range 25,000 to several million. Most RNAs contain a *single* polynucleotide chain, but this can fold back on itself to form double-helical regions containing A : U and G : C base pairs.

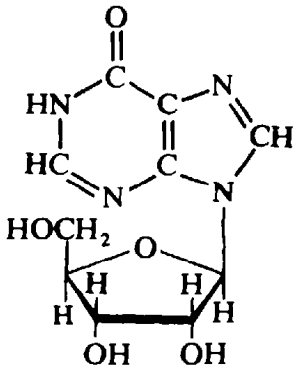
Question: What types of RNA occur in a typical cell and what are their functions?

There are three major types, *transfer RNA* (tRNA), *ribosomal RNA* (rRNA), and *messenger RNA* (mRNA); their roles in the expression of genetic information are treated in detail in Chap. 17.

Transfer RNA ($M_r \cong 25{,}000$) functions as an *adapter* in polypeptide chain synthesis. It comprises 10–20 percent of the total RNA in a cell, and there is at least one type of tRNA for each type of amino acid. Transfer RNAs are unique in that they contain a relatively high proportion of nucleosides of unusual structure (e.g., pseudouridine, inosine, and 2′-*O*-methylnucleosides) and many types of modified bases (e.g., methylated or acetylated adenine, cytosine, guanine, and uracil). As examples, the structures of pseudouridine and inosine are shown below. Inosine has an important role in *codon-anticodon* pairing (Chap. 17).



Pseudouridine
(5-Ribosyluracil)



Inosine

Ribosomal RNA is present in the ribosomes, which contain approximately an equal mass of protein. Ribosomal RNA makes up about 80 percent of the total RNA in the cell and is of several types, distinguished from one another by their sedimentation rates in an ultracentrifuge (Chap. 4). Bacterial ribosomes, for example, contain three types of RNA: 5S, 16S, and 23S. The details of ribosome structure and function are treated in Chap. 17.

Messenger RNA is a very heterogeneous species of RNA. Each molecule carries a copy of a DNA sequence, which is translated in the cytoplasm into one or more polypeptide chains (Chap. 17).

7.10 NUCLEASES

Nucleases are enzymes that degrade nucleic acids by cleaving phosphodiester linkages. They may be specific for DNA or RNA, or they may act on both. A nuclease specific for DNA is called a *deoxyribonuclease* (DNase) and for RNA, *ribonuclease* (RNase).

EXAMPLE 7.18

Is there any specificity with respect to the location of the bond to be cleaved by a nuclease within the polynucleotide chain?

Nucleases are of two general types: (1) *exonucleases* and (2) *endonucleases*. Exonucleases bind to a terminus (5′ or 3′) and remove nucleotides either one or a few at a time. Some require a 5′ terminus and operate in the 5′→3′ direction (5′→3′ exonucleases); others (3′→5′ exonucleases) start at a 3′ terminus and degrade in the opposite direction. There are some exonucleases that will work from either terminus. Exonucleases show no base or sequence specificity.

Endonucleases do not require a terminus and will catalyze the cleavage of a polynucleotide chain at one or more sites. Frequently, they are specific for certain sites (specific base sequences) within the polynucleotide.

Question: Will nucleases cleave both single-stranded and double-stranded nucleic acids?

Many exo- and endonucleases discriminate between these two forms, although some will catalyze the hydrolysis of both. For example, DNase I (bovine pancreas) and DNase II (calf thymus) hydrolyze both forms. The difference between the two enzymes is that DNase I yields 5′-P-terminated oligonucleotides and DNase II yields 3′-P-terminated oligonucleotides. On the other hand, exonuclease III (*E. coli*), a 3′→5′ exonuclease, requires double-stranded DNA as substrate.

Question: What are restriction endonucleases?

Restriction endonucleases are part of a "DNA immunity system" in bacteria. They protect the cell against entry by foreign DNA by catalyzing double-strand cleavages; the cell's own DNA is protected. There are three types of restriction endonucleases. *Type II* restriction endonucleases have been very useful in the analysis and construction of DNA molecules. These enzymes cleave double-stranded DNA at *specific sites* defined by a four- to eight-nucleotide sequence. An important feature of these sequences is their *twofold rotational symmetry*. The sites of cleavage by nearly 300 type II restriction endonucleases have now been determined, and, in many cases, the cleavage is staggered to yield overlapping 3′-hydroxyl or 5′-phosphoryl termini. The recognition sequence and sites of cleavage by the commonly used *Eco*RI endonuclease are shown below. The arrows show the site of cleavage in each strand.

```
      ↓
5′ G A A T T C

C T T A A G 5′
        ↑
```

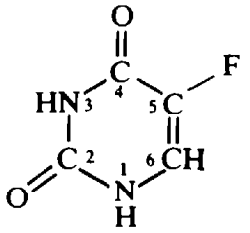
Solved Problems

NUCLEIC ACIDS AND THEIR CHEMICAL CONSTITUENTS; NUCLEOSIDES; NUCLEOTIDES

7.1. Write the structure of 5-fluorouracil.

SOLUTION

Uracil is 2,4-dioxypyrimidine, and 5-fluorouracil is uracil containing a fluorine atom attached to C-5. Thus, its structure is



7.2. 5-Bromouracil is an analog of thymine, but 5-fluorouracil is not. Why?

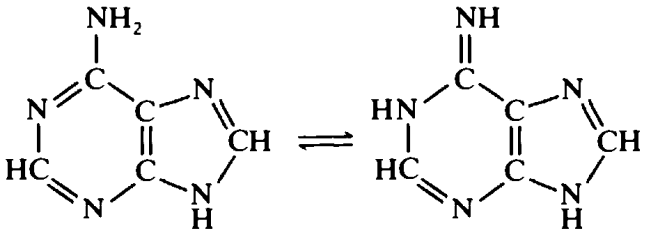
SOLUTION

Thymine is 5-methyluracil. The fluorine atom is much smaller in size than the methyl group, but the bromine atom is similar in size to it. 5-Bromouracil can be readily incorporated into DNA through the action of the enzyme DNA polymerase (Chap. 16), while 5-fluorouracil cannot.

7.3. Write the tautomeric forms of adenine.

SOLUTION

Adenine undergoes amino $\rightleftharpoons$ imino tautomerism. Thus,

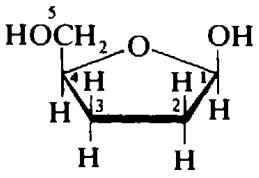


Amino form Imino form

7.4. Write the structure of 2,3-dideoxy-β-D-ribose.

SOLUTION

This compound is β-D-ribose, in which the oxygen atoms are missing from the C-2 and C-3 positions. Thus, its structure is



7.5. If DNA were hydrolyzed so that 2-deoxy-D-ribose were produced, in what anomeric forms would this sugar exist?

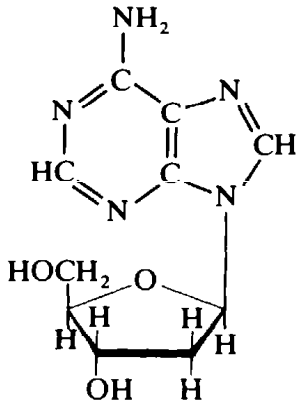
SOLUTION

Both 2-deoxy-β-D-ribose and 2-deoxy-α-D-ribose would be present because the former isomer, which is stabilized within the DNA structure, can now convert to the α form (Chap. 2).

7.6. Write the structure of deoxyadenosine.

SOLUTION

Deoxyadenosine is 9-2′-deoxy-β-D-ribofuranosyladenine, or 2-deoxy-D-ribose linked through a β-glycosidic bond from N-9 of the base to C-1 of the sugar. Thus, its structure is



7.7. Of the nucleosides and nucleoside components adenosine, uridine, D-ribose, deoxyguanosine, cytidine, and thymidine, which would not be expected to occur in significant amounts in a partial hydrolysate of RNA?

SOLUTION

Deoxyguanosine, which contains 2-deoxy-D-ribose (found only in DNA), and thymidine, which contains thymine (present in significant quantities only in DNA), would be absent.

7.8. Match up the compounds, within the following group, that have the same structure: guanosine monophosphate, deoxyguanylic acid, dGDP, GTP, deoxyguanosine monophosphate, GMP, guanosine triphosphate, dGMP, deoxyguanosine diphosphate.

SOLUTION

guanosine monophosphate ≡ GMP

deoxyguanylic acid ≡ dGMP ≡ deoxyguanosine monophosphate

dGDP ≡ deoxyguanosine diphosphate

GTP ≡ guanosine triphosphate

POLYNUCLEOTIDES

7.9. How many 3′,5′ phosphodiester linkages would be present in a linear polynucleotide containing 20 nucleotide units?

SOLUTION

A phosphodiester linkage joins each nucleotide to the adjacent one, so the total number within a polynucleotide is always one less than the number of nucleotide units. Phosphates present at the 5′ or 3′ end of the chain do not constitute phosphodiesters. Therefore, the answer is 19.

7.10. Write the chemical structure of the tetranucleotide ApGpUpCp.

SOLUTION

Remember that sequences are always written in the 5′→3′ direction, from left to right. The 3′ terminus is therefore phosphorylated. The sugar present is D-ribose because there is no *d* prefix to indicate that it is a *deoxy* tetranucleotide. Thus, the structure is as shown below.

7.11. What would be the charge carried by ApGpUpC at neutral pH?

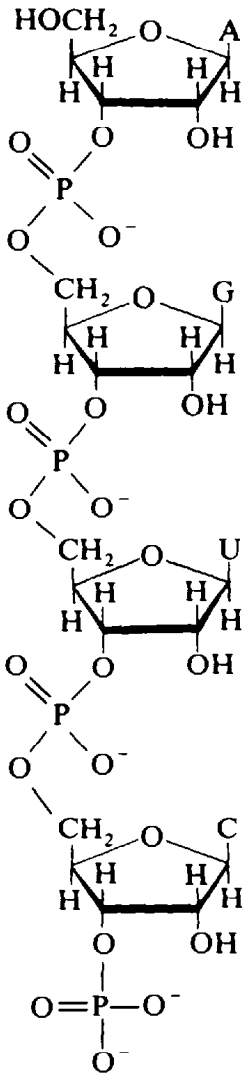
SOLUTION

This tetranucleotide would contain three phosphates, each dissociating one proton. The rest of the molecule would be uncharged at neutral pH. Thus, the charge would be −3.

7.12. Do polynucleotides containing both DNA and RNA within a single covalent structure occur naturally?

SOLUTION

Yes, but in small quantities and only transiently. The nascent (or Okazaki) fragments formed through discontinuous DNA replication contain a short stretch of RNA which serves as a primer for DNA chain growth (Chap. 16).



STRUCTURE OF DNA

7.13. Write the complementary DNA sequence for the following: GCTTAGTA.

SOLUTION

In complementary base pairing within DNA, G pairs with C and A with T. Thus the answer is CGAATCAT.

7.14. Write the complementary RNA sequence for the sequence shown in Prob. 7.13.

SOLUTION

In pairing between DNA and RNA, A and G in DNA link, respectively, with U and C in RNA. Thus the answer is CGAAUCAU.

7.15. With respect to base pairing, 5-methylcytosine behaves in the same way as does cytosine. How is this possible?

SOLUTION

In 5-methylcytosine, the methyl group is positioned away from the region of H bonding with guanine and does not interfere with the formation of the three H bonds occurring in a G:C base pair.

7.16. Why does a fiber of DNA shorten upon drying?

SOLUTION

The DNA duplex exists largely in the B form at high humidities. The pitch (or repeat distance) is 3.4 nm, which covers 10 base pairs. At low humidities, DNA converts to the A form, in which the pitch is less, 2.8 nm, and covers 11 base pairs.

7.17. By approximately what percent would a fiber of DNA shorten upon transfer from an environment of high to low humidity, i.e., following conversion from the B to the A form of DNA?

SOLUTION

Consider 100 base pairs of duplex DNA. In the B form this would have a length of 34 nm (10 base pairs per 3.4 nm). In the A form it would be reduced to approximately 25.5 nm (11 base pairs per 2.8 nm). Thus, the length would shorten by about 25 percent.

7.18. The A and G composition (in mole percent) of one of the strands of a duplex DNA is A = 27 and G = 30. What would be the T and C contents of the complementary strand?

SOLUTION

In duplex DNA, T (mole percent) in one strand equals A in the other, and C equals G. Thus, in the complementary strand T = 27 and C = 30 mole percent.

7.19. With respect to the strand of duplex DNA referred to in Prob. 7.18, what can be said about the A and G contents of its complementary strand?

SOLUTION

From the data provided, knowledge of the individual A and G contents of the complementary strand is not possible. However, A + G would together comprise 100 − (27 + 30), or 43, mole percent of the complementary strand.

DENATURATION OF DNA

7.20. Why does circular duplex DNA renature more rapidly than linear duplex DNA?

SOLUTION

When denatured, the component single strands of circular DNA remain interlocked (assuming neither is broken) and, upon renaturation, can find one another more readily than the completely separated strands of linear DNA.

7.21. Why does DNA denature when it is put into pure water, i.e., at an ionic strength of ~0?

SOLUTION

The melting (or denaturation) temperature of DNA is dependent on ionic strength; lowering the ionic strength lowers the melting temperature. In the extreme situation of zero ionic strength, the electrostatic repulsion within the DNA, whose anionic groups are not shielded by counterions, is sufficient to lower the melting temperature to less than 20°C.

7.22. Why does viral DNA have less sequence heterogeneity than does bacterial DNA?

SOLUTION

The genome of viruses carries fewer genes than does that of bacteria. Thus, in a fixed amount of DNA (say several viral genome equivalents), viral genes would be present in more copies than would individual bacterial genes. Different genes have different nucleotide sequences, but various sequences would be repeated more frequently in the viral DNA; i.e., it would have less sequence heterogeneity.

SIZE, ORGANIZATION, AND TOPOLOGY OF DNA

7.23. The *E. coli* chromosome has a size of approximately 4,000 kb. What length of DNA (B form) would be contained in it?

SOLUTION

The B form of DNA has a length of 3.4 nm per 10 base pairs, or 340 nm per kb. Thus the total length would be 1.36×10^6 nm, or approximately 1.4 mm.

7.24. What length of B-form DNA would be present in a human cell (diploid)?

SOLUTION

The size of the human genome is 2.9×10^6 kb. A diploid cell would contain 5.8×10^6 kb of DNA. This is equivalent to approximately 2×10^9 nm or 2 m of B-form DNA.

7.25. The nucleus of a human cell is of the order of 10 μm in diameter. How can it accommodate 2×10^6 μm of duplex DNA?

SOLUTION

The DNA is condensed, by complexing with histones, into chromatin. The basic unit of condensation is the nucleosome, and these repeating structures (200 base pairs of DNA each) are further organized into more compact structures that make up the chromosomes.

7.26. What features of DNA structure are essential for supercoiling?

SOLUTION

Supercoiling results from the introduction (or removal) of extra double-helical twists that are maintained within the DNA structure. Thus, for supercoiling, the DNA must be double-helical, and it

must be topologically constrained, e.g., by being circular or by being complexed to protein in such a way as to "tie" two strands at certain points.

7.27. If the two strands of a circular DNA molecule were pulled apart (unwound) at a certain position, what type of supercoiling would be introduced into the rest of the molecule?

SOLUTION

In the double-helical section of the molecule that remains, the DNA would be overwound. Thus, extra right-hand turns would be introduced into this section of DNA. This would cause positive supercoiling.

STRUCTURE AND TYPES OF RNA

7.28. What are the main chemical differences between RNA and DNA?

SOLUTION

The sugar present in RNA is ribose, while DNA contains deoxyribose. Also, uracil is present in RNA; this is replaced by thymine in DNA.

7.29. What is the most abundant species of RNA in a typical cell?

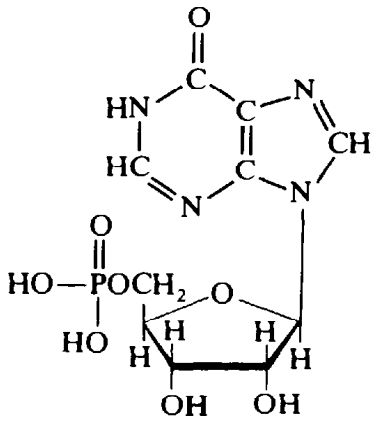
SOLUTION

Ribosomal RNA (rRNA) comprises about 80 percent of all RNA present. It is by far the most abundant.

7.30. Write the structure of inosine 5′-phosphate (IMP).

SOLUTION

IMP consists of inosine to which a phosphate group is attached at C-5 of the ribose residue. Thus, its structure is



7.31. How would the melting curve for RNA compare with that for DNA?

SOLUTION

DNA is nearly always double-helical. Upon heating through the melting temperature, it undergoes a transition from the native (double-helical) to denatured (random-coil) state over a relatively narrow range of temperature. This is accompanied by a 40 percent increase in A_{260} (absorbance at wavelength 260). RNA, on the other hand, is nearly always single-stranded, and the extent of intrastrand base pairing is generally low and variable. Furthermore, the short base-paired (or double-helical) regions vary in

stability. Thus, increasing the temperature of a solution of a typical RNA would result in a *gradual* increase in A_{260}, reflecting the successive melting of short helical regions, and the extent of increase would be significantly less than 40 percent.

NUCLEASES

7.32. What type of bond is cleaved by a nuclease?

SOLUTION

Nucleases cleave 3′,5′ phosphodiester linkages. They always catalyze the hydrolysis of the ester bond where it is connected to either the C-3 or the C-5 of the sugar moiety to give a 5′ or 3′ phosphate plus a 3′ or 5′ OH.

7.33. Why wouldn't an exonuclease degrade the DNA from bacteriophage ϕX174?

SOLUTION

Exonucleases require a 5′ or 3′ terminus of DNA or RNA as substrate. ϕX174 DNA is a single-stranded circle and has no 5′ or 3′ terminus.

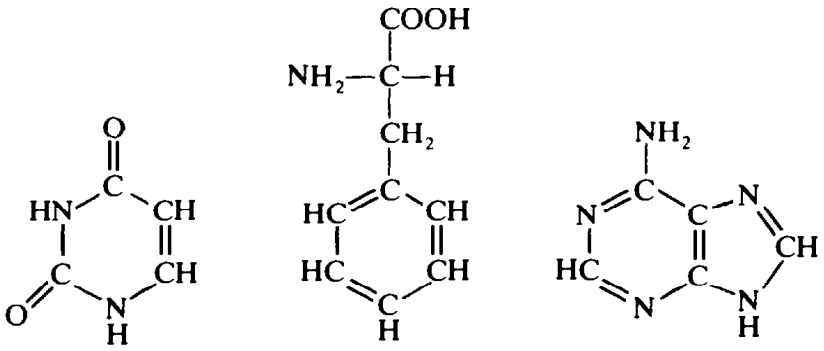
7.34. Approximately how many *Eco*RI sites would you expect to find in the *E. coli* chromosome?

SOLUTION

*Eco*RI recognizes and cleaves at a sequence-specific six-base-pair site on DNA. Considering that each site can be occupied by one of four correctly orientated base pairs, the enzyme would cleave a random sequence once every 4^6 (or 4,096) base pairs. The *E. coli* chromosome has a size of 4.6×10^6 base pairs. Thus, there would be about $4.6 \times 10^6/4{,}096$, or approximately 1,000, *Eco*RI sites on the chromosome.

Supplementary Problems

7.35. Which of the following compounds are pyrimidines or purines?



(*a*) (*b*) (*c*)

7.36. List the common nitrogenous bases found in (*a*) RNA and (*b*) DNA.

7.37. (*a*) Write the structure of 5-methylcytosine. (*b*) Is it a commonly occurring constituent of DNA?

7.38. Write the imino form of cytosine.

7.39. Write the structures of (*a*) deoxythymidine and (*b*) ribothymidine (occurs as a minor constituent in some forms of RNA).

7.40. Give an example of an *N*-glycoside.

7.41. Distinguish between a *nucleoside* and a *nucleotide*.

7.42. Write the structure of 3',5'-AMP (cyclic AMP).

7.43. Which of the following compounds is different from the others: (*a*) GMP, (*b*) deoxyguanosine monophosphate, (*c*) guanylic acid, (*d*) guanosine 5'-phosphate?

7.44. The tetranucleotide d-ApGpUpCp is unlikely to be formed by partial hydrolysis of DNA. Why?

7.45. Would DNA from *Clostridium perfringens* be more or less resistant to denaturation than that from *Alcaligenes faecalis* (see Table 7.1)?

7.46. What is meant by the term *sequence complementarity*?

7.47. List three of the major structural differences between the B and Z forms of DNA.

7.48. Why is the double-helical structure of DNA more stable in solution at higher rather than lower ionic strength?

7.49. Why do nonrepeated sequences of bacterial DNA renature more rapidly than those of eukaryotic DNA?

7.50. What is meant by the term *DNA sequence complexity*?

7.51. Distinguish between the terms *chromosome* and *genome*.

7.52. Distinguish between *core* and *linker* histones.

7.53. What feature of the amino acid composition of histones enables them to interact strongly with DNA?

7.54. (*a*) Distinguish between *positive* and *negative* supercoiling of DNA. (*b*) What is meant by *relaxation* of supercoiled DNA, and (*c*) how can it be achieved?

7.55. List the major differences between the chemical composition of DNA and RNA.

7.56. Why does messenger RNA (mRNA) have the most heterogeneous base sequences of the major forms of RNA present in the cytoplasm of a cell?

7.57. Which of the following sequences would not be cut by a restriction endonuclease and why? (*a*) GAATTC; (*b*) GTATAC; (*c*) GTAATC; (*d*) CAATTG.

Chapter 8

Enzyme Catalysis

8.1 BASIC CONCEPTS

Question: What are enzymes?

Enzymes are proteins that are catalysts of biochemical reactions. They usually exist in very low concentrations in cells, where they increase the *rate* of a reaction without altering its equilibrium position; i.e., both forward *and* reverse reaction rates are enhanced by the same factor. This factor is usually around 10^3–10^{12}.

EXAMPLE 8.1

Although phenomena of fermentation and digestion had long been known, the first clear recognition of an enzyme was made by Payen and Persoz (*Ann. Chim.* (*Phys*), **53**, 73, 1833) when they found that an alcohol precipitate of malt extract contained a thermolabile substance that converted starch into sugar.

EXAMPLE 8.2

The above-mentioned substance was called *diastase* (Greek: "separation") because of its ability to separate soluble dextrin from insoluble envelopes of starch grains. Diastase became a generally applied term for these enzyme mixtures until 1898, when Duclaux suggested the use of *-ase* in the name of an enzyme; this classification procedure still holds today.

EXAMPLE 8.3

Many enzymes were purified from a large number of sources, but it was J. B. Sumner who was the first to *crystallize* one. The enzyme was urease from jack beans. For his travail, which took over 6 years (1924–1930), he was awarded the 1946 Nobel prize. The work demonstrated once and for all that enzymes are distinct chemical entities.

EXAMPLE 8.4

Carbon dioxide gas dissolves readily in water and is spontaneously hydrated to form carbonic acid, which rapidly dissociates to a proton and a bicarbonate ion:

$$CO_2 + H_2O \rightleftharpoons H^+ + HCO_3^-$$

The forward hydration reaction rate for 20 mmol L^{-1} CO_2 at 25°C and pH 7.2 is ~0.6 mmol L^{-1} s^{-1}.

In mammalian red blood cells, the enzyme *carbonic anhydrase* is present at a concentration of 1–2 g per liter of cells; its M_r is 30,000; thus its *molar concentration* is $\sim 50 \times 10^{-6}$. The flux through the forward reaction in the presence of this concentration of enzyme, under the above-mentioned conditions, is ~50 mol L^{-1} s^{-1}, a *rate enhancement* over the noncatalyzed process of 8×10^4.

There are over 2,500 different biochemical reactions with specific enzymes adapted for their rate enhancement. Since different species of organism produce different structural variants of enzymes, the number of different enzyme proteins in all of biology is well in excess of 10^6. Each enzyme is characterized by *specificity* for a narrow range of chemically similar *substrates* (reactants) and also other molecules that modulate their activities; these are called *effectors* and can be *activators*, *inhibitors*, or both; in more complex enzymes, one compound may have either effect, depending on

other physical or chemical conditions. Enzymes range in size from large multiple-subunit complexes (called *multimeric* enzymes; $M_r \simeq 10^6$) to small single-subunit forms.

EXAMPLE 8.5

Aspartate carbamoyltransferase catalyzes the formation of carbamoyl aspartate from carbamoyl phosphate and aspartate in the first committed step of pyrimidine biosynthesis (Chap. 15). The enzyme from the bacterium *E. coli* (M_r = 310,000) consists of 12 subunits, six regulatory and six catalytic. CTP is a negative effector; i.e., it inhibits the enzyme, and does so through binding to the regulatory subunits. ATP is a positive effector that acts through the regulatory subunits, while succinate inhibits the reaction by direct competition with aspartate at the active site (see Chap. 9 for more on effectors).

The surface area of even the smallest enzymes (such as ribonuclease, M_r = 12,000) that is occupied by the chemical groups to which the reactants bind is less than 5 percent of the total area; this region is called the *active site*.

Question: What part of an enzyme is responsible for its substrate specificity?

The particular arrangement of an enzyme's amino acid side chains in the active site determines the type of molecules that can bind *and* react there; there are usually about five such side chains in any particular enzyme. In addition, many enzymes have small nonprotein molecules associated with or near the active site that determine substrate specificity. These molecules are called *cofactors* if they are noncovalently linked to the protein; they are called *prosthetic groups* if covalently bound. In some enzymes a specific metal ion is required for activity.

EXAMPLE 8.6

Carbonic anhydrase has one Zn^{2+} ion per molecule of enzyme, and the metal ion resides in the active site. *Aspartate carbamoyltransferase* has six Zn^{2+} ions per dodecamer; these are required fo the stabilization of the complex, since without Zn^{2+}, the hexamer dissociates.

8.2 CLASSIFICATION OF ENZYMES

Question: On what basis are enzymes given their particular names?

All enzymes are named according to a classification system designed by the Enzyme Commission (EC) of the International Union of Pure and Applied Chemistry (IUPAC) and based on the type of reaction they catalyze. Each enzyme type has a specific, four-integer *EC number* and a complex, but unambiguous, name that obviates confusion about enzymes catalyzing similar but not identical reactions. In practice, many enzymes are known by a *common name*, which is usually derived from the name of its principal, specific reactant, with the suffix *-ase* added. Some common names do not even have *-ase* appended, but these tend to be enzymes studied and named before systematic classification of enzymes was undertaken.

EXAMPLE 8.7

Examples of typical enzyme names are *arginase*, which acts on *arginine*, and *urease*, which acts on urea (Chap. 15). Two atypical common names are *pepsin*, a digestive tract *proteolytic* enzyme (EC number 3.4.23.1), and, more exotically, *rhodanese* (thiosulfate : cyanide sulfurtransferase, EC 2.8.1.1), which is in mammalian liver and kidney and catalyzes the removal of cyanide and thiosulfate from the body. In the latter case, it is understandable why the old name has remained in common use.

The first integer in the EC number designates to which of the six major classes an enzyme belongs (see Table 8.1 for details).

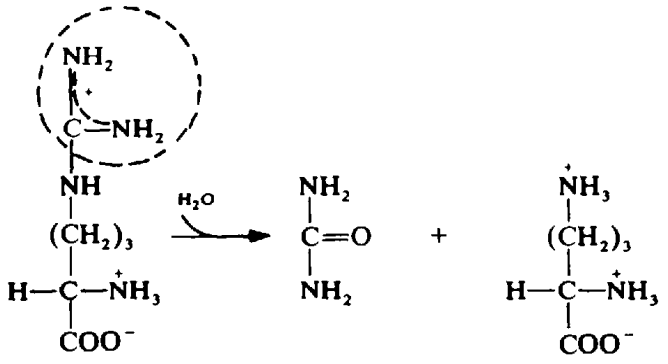
The second integer in an EC enzyme number indicates a sub-class; for hydrolases, this second integer indicates the *type of bond* acted upon by the enzyme (Table 8.2).

Table 8.1. Major Enzyme Classes

First EC Integer	Enzyme Class	Type of Reaction Catalyzed
1.	Oxidoreductase	Oxidation-reduction. A hydrogen, or electron, donor is one of the substrates.
2.	Transferase	Chemical group transfer of the general form $A-X + B \rightarrow A + B-X$.
3.	Hydrolase	Hydrolytic cleavage of C—C, C—N, C—O, and other bonds.
4.	Lyase	Cleavage (*not* hydrolytic) of C—C, C—N, C—O, and other bonds, leaving double bonds; alternatively, addition of groups to a double bond.
5.	Isomerase	Change of geometrical (spatial) arrangement of a molecule.
6.	Ligase	Ligating (joining together) of two molecules, *with* the accompanying hydrolysis of a compound that has a large ΔG for hydrolysis.

EXAMPLE 8.8

Arginase is a *hydrolase* that is in the liver of urea-producing organisms (*ureoteles*). It catalyzes the reaction:



Arginine Urea Ornithine

The official EC name of this enzyme is *L-arginine amidinohydrolase*; the last word refers to the fact that the *amidino group* (dotted circle in the equation) is cleaved from arginine by introduction of a molecule of water across the C—N bond. In the reaction, a nonpeptide C—N bond is cleaved; thus, the second EC number for arginase is 5; its whole classification number is 3.5.3.1.

The third number is a subclassification of the bond type acted upon, or the group transferred in the reaction or both, and the categories vary from one main EC class to the next. The fourth number is simply a serial number.

8.3 MODES OF ENHANCEMENT OF RATES OF BOND CLEAVAGE

The basic mechanisms by which enzymes increase the rates of chemical reactions can be classified into four groups.

Facilitation of Proximity

This effect is also called the *propinquity effect* and means that the rate of a reaction between two molecules is enhanced if they are abstracted from dilute solution and held in close proximity to each other in the enzyme's active site; this raises the *effective* concentration of the reactants.

Table 8.2. Hydrolase Subclassification

First Two EC Integers	Type of Bond Acted Upon
3.1	Ester, $-\overset{\overset{O}{\|}}{C}-O-R$, or with S or P in place of C, or $-\overset{\overset{O}{\|}}{C}-S-R$
3.2	Glycosyl, sugar$-C-O-R$, or with N or S in place of O
3.3	Ether, $R-O-R'$, or with S in place of O
3.4	Peptide, C—N
3.5	Nonpeptide, C—N
3.6	Acid anhydride, $R-\overset{\overset{O}{\|}}{C}-O-\overset{\overset{O}{\|}}{C}-R'$
3.7	C—C
3.8	Halide (X), C—X, or with P in place of C
3.9	P—N
3.10	S—N
3.11	C—P

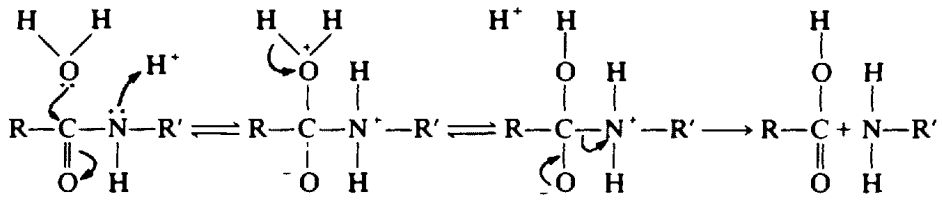
Covalent Catalysis

The side chains of amino acids present a number of *nucleophilic* groups for catalysis; these include $RCOO^-$, $R-NH_2$, aromatic—OH, histidyl, R—OH, and RS^-. These groups attack electrophilic (electron-deficient) parts of substrates to form a covalent bond between the substrate and the enzyme, thus forming a *reaction intermediate*. This type of process is particularly evident in the group-transfer enzymes (EC Class 2; see Table 8.1). In the formation of a covalently bonded intermediate, attack by the enzyme nucleophile (Enz-X in Example 8.10) on the substrate can result in acylation, phosphorylation, or glycosylation of the nucleophile.

Writing the chemical mechanism of a reaction involves describing the rearrangement of electrons as the substrate is converted to the product via some sort of *transition state(s)*. A useful way of depicting the pathway of rearrangement of bonds is by use of curved arrows that indicate the directions of electron flow.

EXAMPLE 8.9

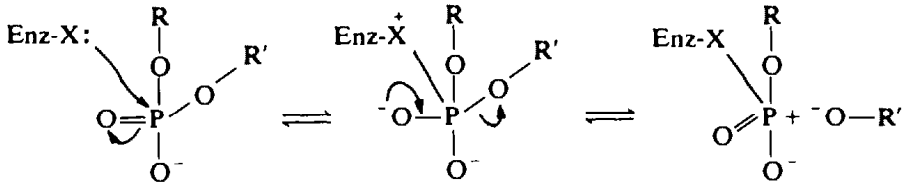
The electron flow diagram for the hydrolysis of a peptide bond is as follows:



The nucleophilic H_2O attacks the electrophilic carbonyl carbon, which is rendered *electron-depleted* by the electron withdrawing attached carbonyl oxygen. Note the tetrahedral intermediates in the second and third structures; in these, the carbon has the usual tetrahedral arrangement of four bonds.

EXAMPLE 8.10

A phosphoenzyme intermediate is formed in one type of covalent catalysis in enzymes:



Numerous examples of this basic mechanism of catalysis can be found among the EC Class 2 enzymes. One example is *hexokinase*.

The covalent intermediates can be attacked by a second nucleophile to cause the release of the product. When the second nucleophile is water, the overall reaction is called *hydrolysis*. Also, in many cases the nucleophile is not simply an amino acid side chain of the enzyme but a prosthetic group; an example is *pyridoxal phosphate* in the *transaminases* (Chap. 15).

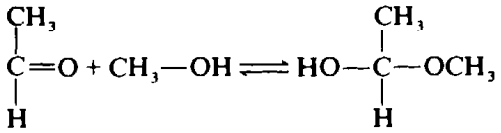
General Acid-Base Catalysis

Acid-base catalysis is defined as the process of *transferral of a proton* in the transition state. It does not involve covalent bond formation per se, but an overall enzymatic reaction can involve this as well.

EXAMPLE 8.11

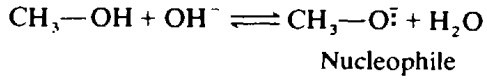
An example of general acid-base catalysis from organic chemistry illustrates the above-mentioned point, but note that hemiacetals also form in some enzymatic reactions.

Overall reaction:

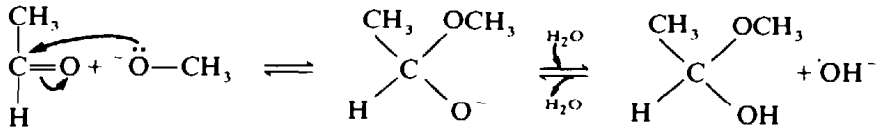


Acetaldehyde Methanol Hemiacetal

Reaction mechanism A: A *base* (OH^-) accelerates hemiacetal formation as follows:



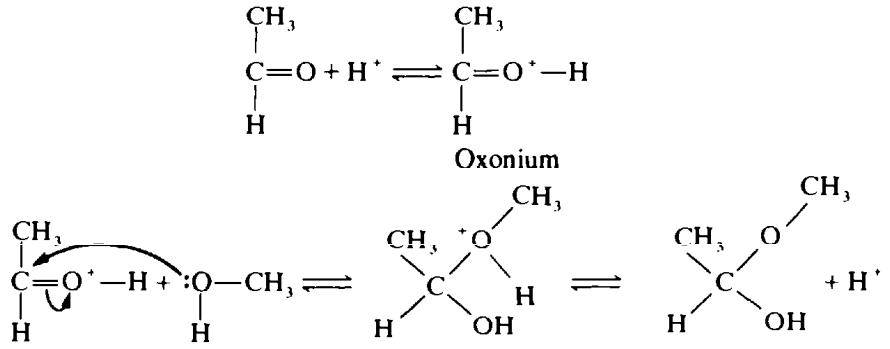
Nucleophile



Nucleophilic attack

Note: The OH^- is recycled in the reaction; thus, it can be considered to be a catalyst in the true sense of the word.

Reaction mechanism B: Acid catalysis also occurs in the reaction, and it involves the formation of the *oxonium* salt, followed by reaction with the alcohol as follows:



In the preceding example, the rate of hemiacetal formation is enhanced in strong acid *or* strong base. In other cases, only one, either base or acid, might be a catalyst.

EXAMPLE 8.12

The hydrolysis of *nitramide* is susceptible to base, but not acid, catalysis. An elevation of pH leads to an increased rate of reaction with no net consumption of base, as is shown in the following reaction:

$$NH_2NO_2 + OH^- \rightarrow H_2O + NHNO_2^-$$
$$NHNO_2^- \rightarrow N_2O + OH^-$$

OH^- is not the only base that will catalyze the hydrolysis; other bases such as acetate also react; e.g.,

$$NH_2NO_2 + CH_3COO^- \rightarrow CH_3COOH + NHNO_2^-$$
$$NHNO_2^- \rightarrow N_2O + OH^-$$
$$OH^- + CH_3COOH \rightarrow H_2O + CH_3COO^-$$

According to the Brønsted-Lowry definitions, and as implied in the previous example, an acid is any moiety that will donate a proton while a base is one that will accept a proton from another moiety.

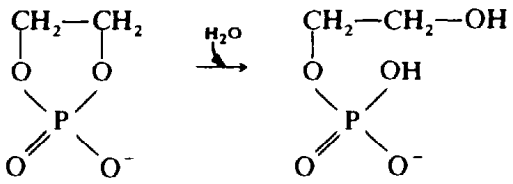
Acid-base catalysis does not contribute to rate enhancement by a factor greater than ~100, but together with other mechanisms that operate in the active site of an enzyme, it contributes considerably to increasing the enzymatic rate of reactions. The amino acid side chains of glutamic acid, histidine, aspartic acid, lysine, tyrosine, and cysteine in their *protonated* forms can act as *acid catalysts* and in their *unprotonated* forms as base catalysts (see Prob. 8.11). Clearly, the effectiveness of the side chain as a catalyst will depend on the pK_a (Chap. 3) in the environment of the active site and on the pH at which the enzyme operates.

Strain, Molecular Distortion, and Shape Change

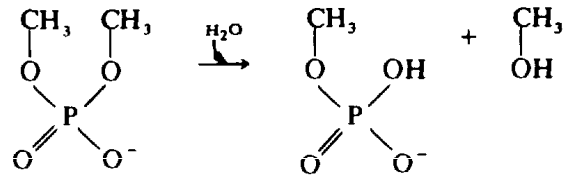
Strain in the bond system of reactants and the release of the strain as the transition state converts into the products (like cutting a wound clock spring) can provide rate enhancement of chemical reactions.

EXAMPLE 8.13

The following two chemical reactions involve hydrolysis of a phosphate ester bond.



(*a*)



(*b*)

Under standard conditions, reaction (*a*) is 10^8 times faster than reaction (*b*). The explanation is that the cyclic compound in (*a*) has considerable *bond strain* (potential energy in this configuration is high), which is released on ring opening during hydrolysis. This type of strain is not present in the diester in (*b*).

In the case of enzymes, not only may the substrate be distorted (have strain) but an extra degree of freedom is introduced, namely, the enzyme with all its amino acid side chains. The binding of a substrate to an enzyme involves *interaction energy*, which may facilitate catalysis. Also for an increase in catalytic rate, there must be an overall *destabilization* of the enzyme-substrate complex and an increase in the stability of the transition state. This idea is illustrated in Fig. 8-1.

In the *uncatalyzed* reaction [Fig. 8-1(*a*)], the reactant has a relatively low probability of assuming the *strained* conformation necessary for interaction between the two reactive groups. In order for the reaction to take place, the molecule must cross this so-called *activation-energy barrier*. In the *catalyzed* reaction [Fig. 8-1(*b*)], the *binding* of the reactant to the enzyme leads to the formation of a combined structure (enzyme-substrate complex) in which the tendency for the substrate to form into the *transition state* is greater; i.e., less energy is involved in bringing the reactive groups together. Therefore, the reaction proceeds faster.

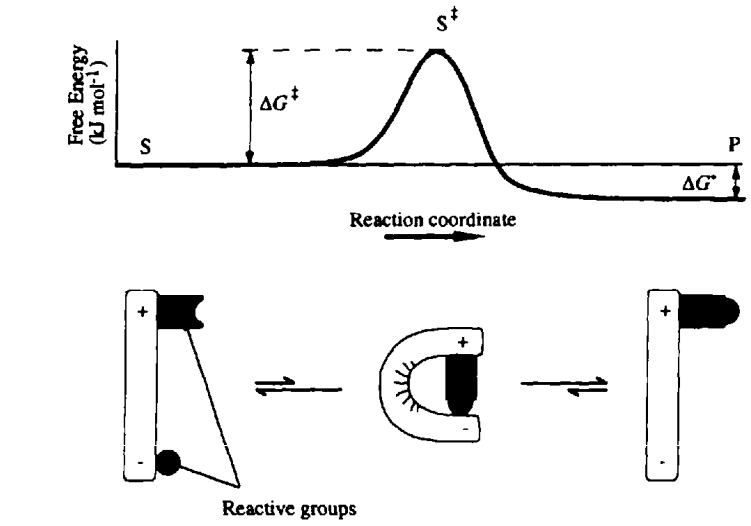
The destabilization of the enzyme-substrate complex can be imagined to be due to distortion of bond angles and lengths from their previously more stable configuration; this may be achieved by electrostatic attraction or repulsion by groups on the substrate and enzyme. Or, it could involve *desolvation* (removal of water) of a charged group in a hydrophobic active site. A further consideration is that of *entropy* change in the reaction; this is discussed in the next section.

Question: Does tight binding between an enzyme and its substrate imply rapid catalysis?

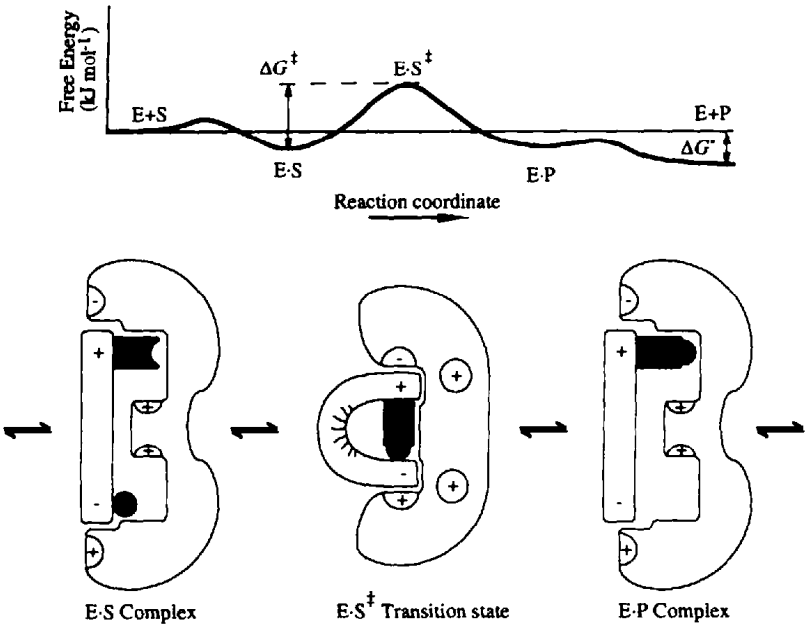
If a substrate were to bind *without* significant transformation of binding energy into distortion strain, then binding would be stronger. But this would not necessarily affect $\Delta G^{\ddagger}$ (see Fig. 8-1). However, if some of the free energy of binding were to be used to distort the substrate more toward the shape of the transition state, or to distort the enzyme to be more *complementary* to the shape of the transition state, then binding of the enzyme to the substrate would be weakened, while binding to the transition state of the substrate would be correspondingly enhanced. Hence, to dispel a common belief, tight binding of the *substrate* is not necessarily useful in enhancing the rate of an enzyme reaction.

EXAMPLE 8.14

Suppose a substrate half-saturates the active sites of a solution of enzyme when present at $10^{-7}\,\text{mol}\,\text{L}^{-1}$ (i.e., $K_d = 10^{-7}\,\text{mol}\,\text{L}^{-1}$), but the concentration under *physiological* conditions is $10^{-3}\,\text{mol}\,\text{L}^{-1}$. Under



(a) Uncatalyzed reaction



(b) Catalyzed reaction

Fig. 8-1 Activation energy is lowered in catalyzed reactions. The graphs above each reaction scheme indicate the energy of the substrate (depicted here as *potential* energy of the *bent* substrate) at each stage of the reaction. The arrows indicate, according to their length and boldness, the probability and, in this case, the rate of the reactions. $\Delta G^{\ddagger}$ is the activation energy of the transition-state(s) molecule, and ΔG° is the *overall* free energy of the reaction (Chap. 10). N.B. Changes in the enzyme and substrate lead to binding being tighter in the transition state than in either the E.S or E.P states.

physiological conditions, the enzyme sites are fully saturated (i.e., all sites are filled), so the enzyme rate enhancement is not what it could be if a large portion of the binding energy were used for *destabilizing* the enzyme-substrate (E.S) complex.

If some of the binding energy were used to introduce strain or distortion within the enzyme or substrate molecules, such that tighter binding of the transition state were achieved, then the binding affinity of the enzyme for the substrate would be reduced.

Many enzymes, in fact, have binding affinities for their substrates that have values that are around the mean physiological concentrations, possibly as a result of evolutionary pressure for efficient catalysis.

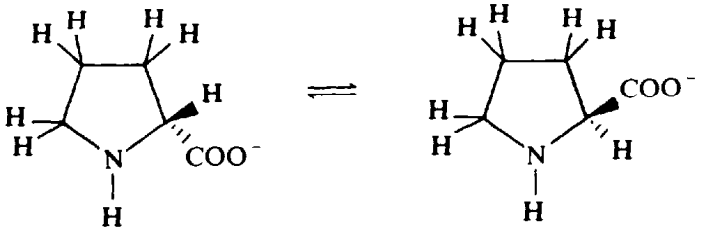
EXAMPLE 8.15

x-Ray analysis of crystals of *carboxypeptidase A* (a pancreatic exopeptidase) with bound *pseudo substrate* (a false substrate that is not degraded by the enzyme, i.e., an inhibitor), indicates that the *susceptible* peptide bond is twisted out of the *normal* planar configuration usually seen in peptide bonds (Chap. 4). This distortion leads to a loss of resonance energy in the bond, which enhances its susceptibility to hydrolytic attack.

Because in catalysis the enzyme-substrate complex is destabilized and the energy so involved is released on forming the transition state, the enzyme binds the substrate very tightly in the transition state. Some enzymes can be dramatically inhibited by so-called *transition-state analogs*. The transition state normally has only a fleeting existence ($<10^{-13}$ s), but the analogs are stable structures that *resemble* the postulated transition-state complex.

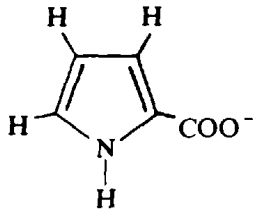
EXAMPLE 8.16

Proline racemase, a bacterial enzyme, catalyzes the interconvention of the D and L isomers of proline:



L-Proline D-Proline

It was postulated that in proceeding from the L to the D isomer, a *planar* (rather than the usual tetrahedral) configuration of the molecule momentarily exists at the α carbon.



Pyrrole 2-carboxylate

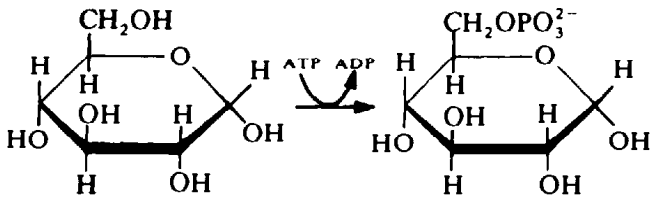
A planar analog of proline is *pyrrole 2-carboxylate*, and this proves to be a potent inhibitor of the racemase; it gives rise to 50 percent inhibition at a concentration 160 times less than the concentration of D- or L-proline that gives 50 percent binding. Thus, it is a good example of a transition-state analog.

Question: Do *both* the enzyme and substrate undergo a change when they interact?

Yes, the concept of *induced fit* of an active site to a substrate emphasizes the *adaptation* of the active site to fit the functional groups of the substrate. A *poor* substrate or inhibitor does not induce the correct conformational response in the active site.

EXAMPLE 8.17

Hexokinase demonstrates the *induced fit* phenomenon; it catalyzes phosphoryl transfer from ATP to C-6 of glucose as follows:



Glucose Glucose 6-phosphate

The enzyme can also catalyze the transfer of the terminal phosphoryl of ATP to water; i.e., it acts as an ATPase but at a rate 5×10^6 times slower than the above reaction. The basic and nucleophilic properties of water versus the C-6 hydroxyl of glucose are sufficiently similar to suggest no marked differences in rate. Therefore, the explanation of the rate difference is that glucose induces a conformational change that *establishes* the correct active-site geometry in the enzyme, whereas a water molecule is too small to do so.

The induction of the correct geometry in the active site of an enzyme is *paid for* by a *good substrate*, with binding energy. An alternative explanation to that of induced fit is that some small molecules (e.g., H_2O in the hexokinase example) bind *nonproductively*, i.e., their small size allows them to assume many orientations with respect to the other substrate (ATP in the case of hexokinase) that do not lead to reaction. Large substrates are restricted in motion and are held in a catalytically *correct* orientation millions of times more often during molecular vibrations than is, say, water.

8.4 RATE ENHANCEMENT AND ACTIVATION ENERGY

Question: Some biochemicals are stable when in pure form on a shelf and yet in the presence of an enzyme break down rapidly. Why?

There is an important distinction to be made between *thermodynamic stability* (expressed in terms of the equilibrium constant of the reaction) and the *kinetic stability* of a substance; the latter merely refers to how fast the reaction proceeds, the former to the final position of the reaction in terms of the relative amounts of substrate and product at equilibrium. (See Example 8.18.) Enzymes affect the kinetic stability of a substance.

EXAMPLE 8.18

Most reduced organic molecules, such as glucose, are *thermodynamically unstable* in our oxidizing atmosphere.

$$\text{glucose} + 6O_2 \rightarrow 6CO_2 + 6H_2O \qquad -\Delta G^\circ = 2{,}872 \text{ kJ mol}^{-1}$$

Thus, oxidation is very *exergonic* (heat-producing), and the reaction is favored by the large $-\Delta G^\circ$ (Chap. 10) of the reaction. But we are all aware that glucose is stable *on the shelf*. Thus, it is *thermodynamically unstable* but *kinetically stable*.

The distinction between kinetic and thermodynamic stability is important and is explained by the concept of the *free energy* of activation necessary to convert the substrate to its transition state. In order for the substrate to form products, its internal free energy must exceed a certain value; i.e., it must *surmount an energy barrier*. The energy barrier is that of the free energy of the transition state, $\Delta G^\ddagger$. The transition-state theory of reaction rates introduced by H. Eyring relates the rate of the reaction to the magnitude of $\Delta G^\ddagger$.

Question: Is there a simple mathematical relationship between reaction rate and $\Delta G^\ddagger$?

Yes. In the 1880s, Arrhenius observed that the rate constant k for a simple chemical reaction varies with the temperature according to

$$k = Ae^{-E_a/RT} \tag{8.1}$$

where E_a is the so-called *Arrhenius activation energy* of the reaction, A is called the *preexponential factor*, R is the universal gas constant, and T is the temperature (K). However, it became apparent that A was *not quite* temperature-independent, especially in catalyzed reactions; thus Eyring proposed that *all* transition states break down with the same rate constant, $\kappa T/h$; where κ and h are Boltzmann's and Planck's constants, respectively. He therefore proposed that for any reaction,

$$k = \frac{\kappa T}{h} e^{-\Delta G^{\ddagger}/RT} \tag{8.2}$$

where, again, $\Delta G^{\ddagger}$ is the activation energy of the *transition-state* complex.

In Chap. 10 it will be shown that the Gibbs free energy of a system is made up of two components such that

$$\Delta G = \Delta H - T\Delta S \tag{8.3}$$

where ΔS is the *entropy* change and ΔH the *enthalpy* change in the reaction system. Therefore, Eq. (8.2) can be written as

$$k = \left(\frac{\kappa T}{h} e^{\Delta S^{\ddagger}/R}\right) e^{-\Delta H^{\ddagger}/RT} \tag{8.4}$$

Entropy is an equilibrium thermodynamic entity that is *interpreted* mechanically as the *degree of disorder* in a system. From Eq. (8.4), it is therefore seen that (1) the preexponential factor A [Eq. (8.1)] can be interpreted as being related to the *organization* of a reactant in an enzyme as the transition-state complex is formed and (2) the exponential factor relates to the enthalpy (heat) of the reaction.

Any molecular factors that tend to stabilize the transition state decrease $\Delta G^{\ddagger}$ and thus increase the rate of the reaction. Thus, this rate enhancement can result from either entropy or enthalpy effects, or from both.

8.5 SITE-DIRECTED MUTAGENESIS

Valuable insights into the mechanism of action of enzymes have been obtained by comparative studies of enzymes from different sources; e.g., different isoenzymes, or enzymes from different species. The amount of variation available from this source, however, is often limited. Further modification of enzymatic activity has been discovered through the study of the effects on enzymatic activity of specific chemical modifications. With the availability of methods for cloning and expressing DNA (Chap. 17), methods have been developed for the specific modification of enzymes through *site-directed mutagenesis*; i.e., alteration of a gene to yield a protein with specific alterations in the sequence. Comparative studies of such mutant enzymes with one another, and with the *wild type* (or naturally occurring form), have allowed valuable information to be gained on the roles of specific amino acid residues in the processes of substrate binding, catalysis, enzyme stability and regulation.

EXAMPLE 8.19

The specificity of trypsin toward peptide bonds adjacent to lysine and arginine has been altered by replacing a glycine residue in the enzyme active site with an alanine, which carries an additional methyl group. This replacement favors the binding of lysine over the somewhat bulkier arginine.

EXAMPLE 8.20

In the catalytic mechanism of the serine protease subtilisin, the tetrahedral intermediate is believed to be stabilized by a hydrogen bond to the side chain of Asn 155. Replacement of Asn 155 with Gly left the substrate binding unaffected, but inhibited the catalytic step, confirming the proposed mechanism.

EXAMPLE 8.21

On the basis of chemical modification studies, Tyr 198 of carboxypeptidase A was proposed to act as a proton donor (i.e., a general acid) in the mechanism of catalysis. However, when Tyr 198 was replaced with Phe by means of site-directed mutagenesis, the modified enzyme retained substantial enzymatic activity, indicating that the tyrosyl hydroxyl may not have a specific role in catalysis.

Question: Does the abolition of activity following mutagenesis unequivocally indicate a catalytic role for the mutated residue?

Not necessarily! It is possible that replacement of one amino acid with another may abolish specific interactions critical for local folding or even overall protein stability. It is essential in these types of study to monitor protein conformation after mutagenesis. It is even more helpful if detailed structures, either from x-ray crystallography or NMR (Chap. 4), are available for the wild-type and mutant forms of the enzymes.

Solved Problems

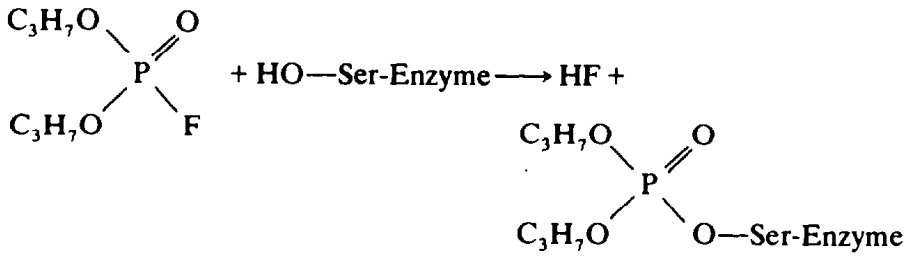
BASIC CONCEPTS

8.1. Justify the claim that the surface area of the active site of an enzyme is less than 5 percent of its total surface area. Consider the specific case of a 27,000-Da globular enzyme with five amino acids in its active site.

SOLUTION

The ratio of the volume of the active site to that of the whole protein is ~5:27,000/110 = 0.02; this is so because the mean amino acid residue weight is ~110. If we assume that the enzyme is spherical, the above volume ratio corresponds to a surface-area ratio of $(0.02)^{2/3} \times 0.5$. The factor 0.5 accounts for the fact that half of the active-site residues face outward and contribute also to the total surface area; the other half faces in toward the interior of the enzyme. The answer is 0.04, or 4 percent of the surface area.

8.2. The nerve gas *diisopropyl fluorophosphate* (DFP) reacts with the serine —OH in some enzymes to form HF and the *O*-phosphoryl ester as follows:



In a chemical enzyme-modification experiment conducted by E. F. Jansen and colleagues in 1949, chymotrypsin was incubated with ^{32}P-labeled DFP and then hydrolyzed with a strong acid. Separation of the constituent amino acids revealed 1 mol of *labeled O*-phosphorylserine per 25,000 g of chymotrypsin. Since DFP is a potent inhibitor of the enzyme *chymotrypsin*, what might we infer about the amino acid side-chain composition of the active site?

SOLUTION

Since for chymotrypsin $M_r = 25{,}000$, only a single serine had reacted out of a total of 27. This indicated that a particular serine is an important component of the active site. This experiment is the archetypal form of many enzyme-modification procedures that are now used frequently to identify active-site constituents.

CLASSIFICATION OF ENZYMES

8.3. What is the order of abundance of enzymes in the six EC groups?

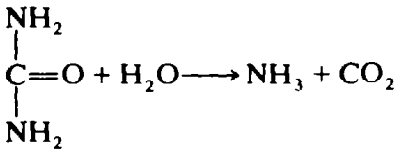
SOLUTION

Of the ~2,500 different, named enzymes, the most abundant group is the oxidoreductases, group 1. The order is 1>2>4>3>6>5.

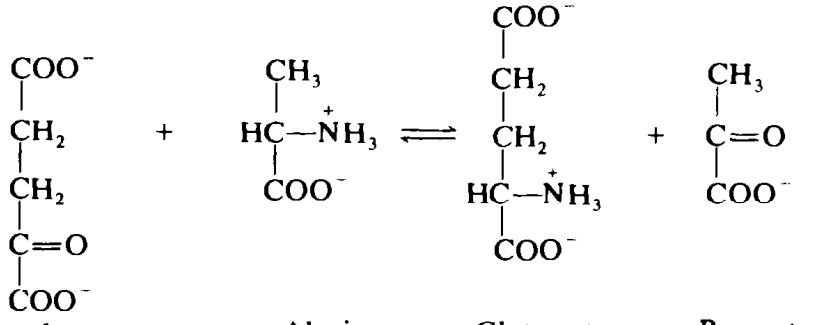
8.4. Classify the following enzyme-catalyzed reactions into their major EC groups and suggest possible common names for each.

(*a*) D-Glyceraldehyde 3-phosphate + P_i + NAD^+ $\rightleftharpoons$ 1,3-bisphosphoglycerate + NADH

(*b*)

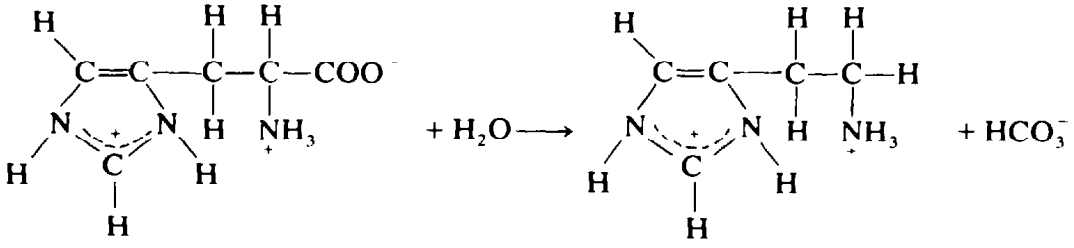


(*c*)



2-Oxoglutarate L-Alanine L-Glutamate Pyruvate

(*d*)



L-Histidine Histamine

(*e*) L-Alanine $\rightleftharpoons$ D-Alanine

(*f*) L-Ribulose 5-phosphate $\rightleftharpoons$ D-Xylulose 5-phosphate

(*g*) The enzyme that catalyzes the rearrangement of S—S bonds in proteins.

(*h*) ATP + L-Tyrosine + $tRNA^{Tyr}$ → AMP + PP_i + L-Tyrosyl-$tRNA^{Tyr}$

(*i*) ATP + γ-L-Glutamyl-L-cysteine + Glycine → ADP + P_i + Glutathione

SOLUTION

(*a*) Glyceraldehyde-3-phosphate dehydrogenase, EC 1.2.1.12. The systematic name is *D-glyceraldehyde-3-phosphate:NAD$^+$ oxidoreductase* (phosphorylating); it is an important glycolytic enzyme.

(*b*) Urease, EC 3.5.1.5. Its systematic name is *urea amidohydrolase*. Interestingly, it is a nickel containing enzyme.

(*c*) Alanine transaminase, EC 2.6.1.12. The systematic name is *L-alanine:2-oxoglutarate aminotransferase*. Note also that aminotransferases almost invariably have *pyridoxal phosphate* as a cofactor.

(*d*) Histidine decarboxylase, EC 4.1.1.22. The systematic name is *L-histidine carboxy-lyase*; it, too, requires pyridoxal phosphate in animals, but the bacterial enzyme does not.

(*e*) Alanine racemase, EC 5.1.1.1. The systematic name is the same. It has the honor of being the first enzyme in group 5; it also requires pyridoxal phosphate as a cofactor. (A *racemic* mixture is a mixture of optical isomers of one chemical species, Chap. 2).

(*f*) Ribulose phosphate epimerase, EC 5.1.3.4. Its official name is *L-ribulose-5-phosphate 4-epimerase*. This is a key enzyme in the pentose phosphate pathway. (An *epimer* is a stereoisomer of a sugar differing in the configuration on only *one* carbon, Chap. 2).

(*g*) Disulfide bond (S—S) rearrangease, EC 5.3.4.1. The official name is *protein disulfide-isomerase*. The common name is an example of naming the enzyme after the phenomenon with which it is associated; enzymes in this category are called *phenomenases*, and the EC of IUPAC and the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB) have generally disapproved of such naming. Another common example is the use of the name *translocase* for an enzyme or protein carrier that catalyzes the movement of a moiety between biological structural compartments, e.g., *ATP translocase* in mitochondria (Chap. 14).

(*h*) Tyrosyl-tRNA synthetase, EC 6.1.1.1. The official name is *L-tyrosine:tRNA^{Tyr} ligase (AMP-forming)*. It is the first of the group 6 enzymes but, more importantly, is essential to life because of its role in protein synthesis (Chap. 17).

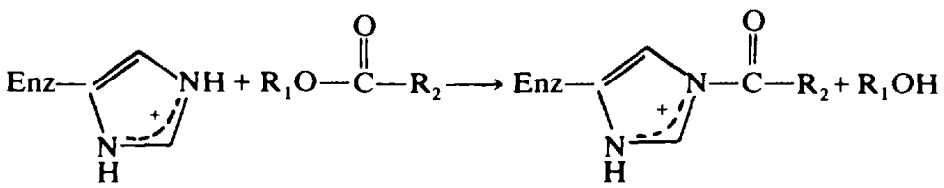
(*i*) Glutathione synthetase, EC 6.3.2.3. The official name is *γ-L-glutamyl-L-cysteine:glycine ligase (ADP-forming)*. This is the key enzyme in glutathione production and is found in many tissues and organisms.

MODES OF ENHANCEMENT OF RATES OF BOND CLEAVAGE

8.5. Covalent enzyme catalysis involves the formation of a transient covalent bond between an enzyme and its substrate. Below are the general structures of commonly encountered so-called *acyl-enzyme intermediates* and other covalent derivatives.

1. $$\text{Enz—OH} + \text{R}_1\text{O—}\overset{\overset{\text{O}}{\|}}{\text{C}}\text{—R}_2 \longrightarrow \text{Enz—O—}\overset{\overset{\text{O}}{\|}}{\text{C}}\text{—R}_2 + \text{R}_1\text{OH}$$

2. $$\text{Enz—SH} + \text{R}_1\text{O—}\overset{\overset{\text{O}}{\|}}{\text{C}}\text{—R}_2 \longrightarrow \text{Enz—S—}\overset{\overset{\text{O}}{\|}}{\text{C}}\text{—R}_2 + \text{R}_1\text{OH}$$

3. 

4. $\text{Enz—NH}_2 + \text{O=C}\begin{matrix}\diagup \text{R}_1 \\ \diagdown \text{R}_2\end{matrix} \longrightarrow \text{Enz—N=C}\begin{matrix}\diagup \text{R}_1 \\ \diagdown \text{R}_2\end{matrix} + \text{H}_2\text{O}$

(*a*) Give examples of amino acid residues that have side chains with the reactive groups indicated above. (*b*) Give chemical names to the acyl-enzyme intermediate compounds.

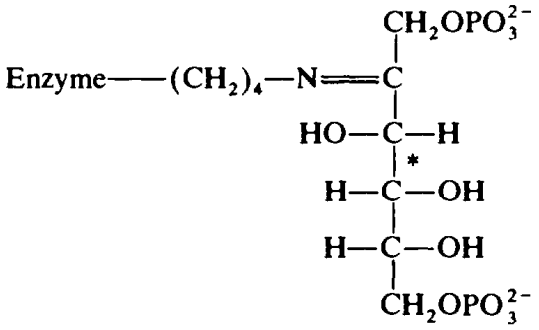
SOLUTION

(*a*) (1) Serine, (2) cysteine, (3) histidine, (4) lysine.
(*b*) (1) Ester, (2) thioester, (3) acylimidazole, (4) Schiff base.

8.6. The glycolytic pathway enzyme *fructose-1,6-bisphosphate aldolase* forms an acyl-enzyme intermediate with its ketone substrate *fructose 1,6-bisphosphate*. Given that the enzyme contains a lysine residue that is essential for its activity, what type of covalent intermediate is likely to be formed?

SOLUTION

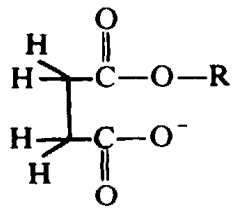
The general structure is that of (4) in Prob. 8.5(*b*), namely, a Schiff base (also called a *ketimine*).



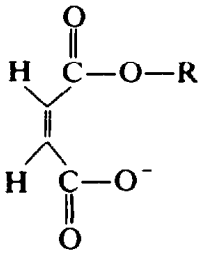
The formation of the ketimine labilizes the bond marked with the asterisk, leading to its cleavage.

8.7. *Propinquity (proximity) effects* are important in reaction rate enhancement. In the case of the following compounds, *anhydrides* (products formed on removal of water) form at different rates. Arrange the compounds in order of their rates of anhydride formation and explain the reasons for the ordering.

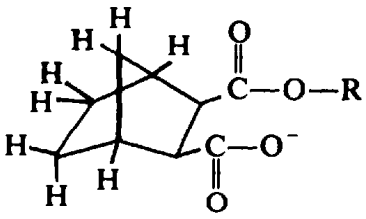
(*a*)



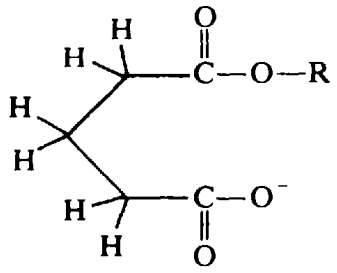
(*b*)



(*c*)



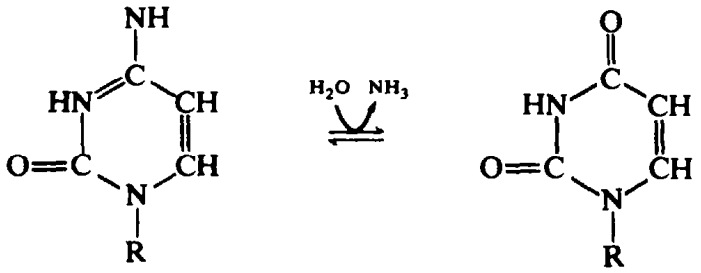
(*d*)



SOLUTION

The relative rates of anhydride formation are as follows: (*d*) 1; (*a*) 230; (*b*) 10,100; (*c*) 53,000. A greater rate enhancement occurs in the compounds in which the reacting carboxyl groups are held more rigidly; this increases the time during which the transition state can form and therefore the time in which the products can be formed.

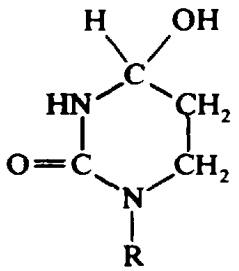
8.8. Transition-state analogs are potent inhibitors of enzymes. In the enzyme *cytidine deaminase* from the bacterium *E. coli*, the following chemical transformation takes place:



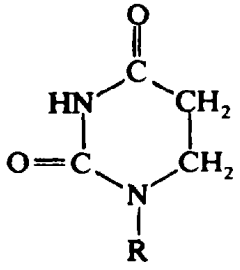
where R denotes a ribose residue.

(*a*) Draw a possible transition-state compound.

(*b*) The two following compounds have different effects on the reaction rate; one is a transition-state analog, while the other is a substrate. Give reasons for your proposal for which is the analog.



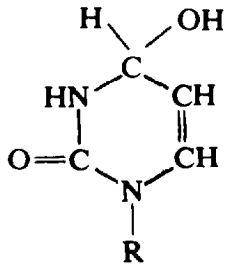
3,4,5,6-Tetrahydrouridine
(1)



5,6-Dihydrouridine
(2)

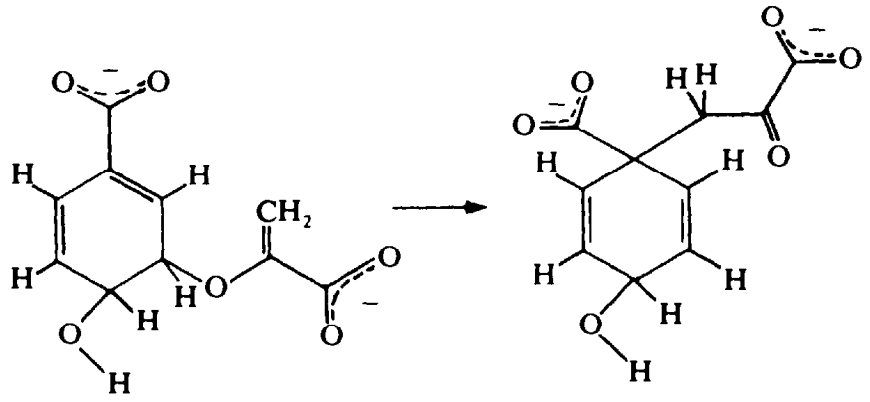
SOLUTION

(*a*) The likely transition state is the so-called *tetrahedral intermediate*; namely, an intermediate species in which the carbon has the usual four-bond arrangement for carbon:



(*b*) It is clear that compound (1) has a structure very similar to that of the intermediate in (*a*); it is indeed a potent inhibitor (transition-state analog) of cytidine deaminase.

8.9. The bacterial enzyme *chorismate mutase–prephenate dehydrogenase* is peculiar because it is a single protein unit with *two* catalytic activities. It catalyzes the sequential reactions of mutation of chorismate to prephenate and then the reaction that leads to the formation of *phenylalanine* and *tyrosine*, through oxidation of prephenate. The first of these reactions is interesting because it is one of the few strictly *single-substrate* enzymatic reactions; it entails the migration of a side chain from one part of the ring to another, as shown in the scheme below.



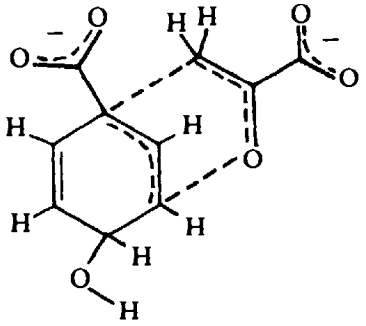
Chorismate Prephenate

(*a*) Predict a likely transition-state structure.

(*b*) Suggest a likely transition-state analog that might be a potent inhibitor of the enzyme.

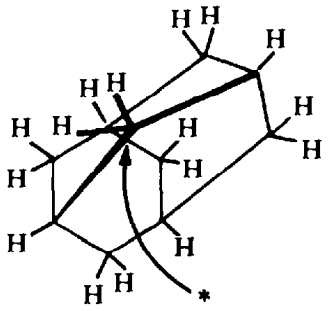
SOLUTION

(*a*) By using molecular orbital calculations, P. R. Andrews and G. D. Smith in 1973 suggested the following structure as that of the transition-state molecule:



Transition state

(*b*) Andrews and Smith recognized the similarity between the structure of the transition state and *adamantane*.



Adamantane

Adamantane has an *extra* methylene bridge (the asterisk in the diagram) linked to the six-membered ring, thus stabilizing a cagelike structure. The authors indeed subsequently showed that some adamantane derivatives are potent inhibitors of chorismate mutase; thus, these are examples of transition-state analogs.

Note: Since the enzyme is not found in mammals, inhibitors of this enzyme *may* be an effective means of controlling bacterial infection. Certainly, *species-selective toxicity* is an important consideration in the development of new antimicrobial agents.

8.10. *Lysozyme* is an enzyme found in tears. It hydrolyzes bacterial cell wall polysaccharides, and it has one of the best understood of all enzyme mechanisms. The enzyme is a single polypeptide chain of 129 amino acids folded into a shape like a grain of puffed wheat, with a cleft along one side. Into the cleft fits the substrate, a polysaccharide made up of alternating units of *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM). Details of the binding of a competitive inhibitor $(NAG)_3$ to the active site have been obtained using x-ray crystallography. Using the x-ray structure, insight into the binding of substrates such as $(NAG\text{-}NAM)_3$ has been obtained (see Fig. 8-2 for a schematic model of the active site and the binding groups).

The enzyme catalyzes the cleavage of the bond between carbon 1 of residue [4] and the oxygen atom of the glycosidic linkage of residue [5]. Two amino acid side chains in the region of this bond can serve as proton donors or acceptors: Asp 52 and Glu 35, each of which is about 0.3 nm from the bond. Asp 52 is in a polar environment and is ionized at the pH optimum of lysozyme (pH 5), whereas Glu 35 is in a nonpolar region and is not ionized. The proposed catalytic mechanism is given in Fig. 8-3.

(*a*) To which EC group does lysozyme belong, and what are the first two integers in its EC number?

(*b*) Describe in words the various basic chemical processes that take place in the cleavage of $(NAM\text{-}NAG)_3$, as shown in Fig. 8-3.

(*c*) What type of catalysis is operating here, covalent or noncovalent?

(*d*) What type of bonds are involved in the binding of the substrate $(NAM\text{-}NAG)_3$ to the enzyme?

(*e*) On binding to the enzyme, the sugar residue [4] is distorted from a *chair* conformation to that of a *half-chair* (Chap. 2). How might this aid catalysis?

SOLUTION

(*a*) Lysozyme is a *hydrolase*; thus its first EC number (Table 8.1) is 3. Since it catalyzes the hydrolysis of a C—O bond, its second number is 2 (Table 8.2).

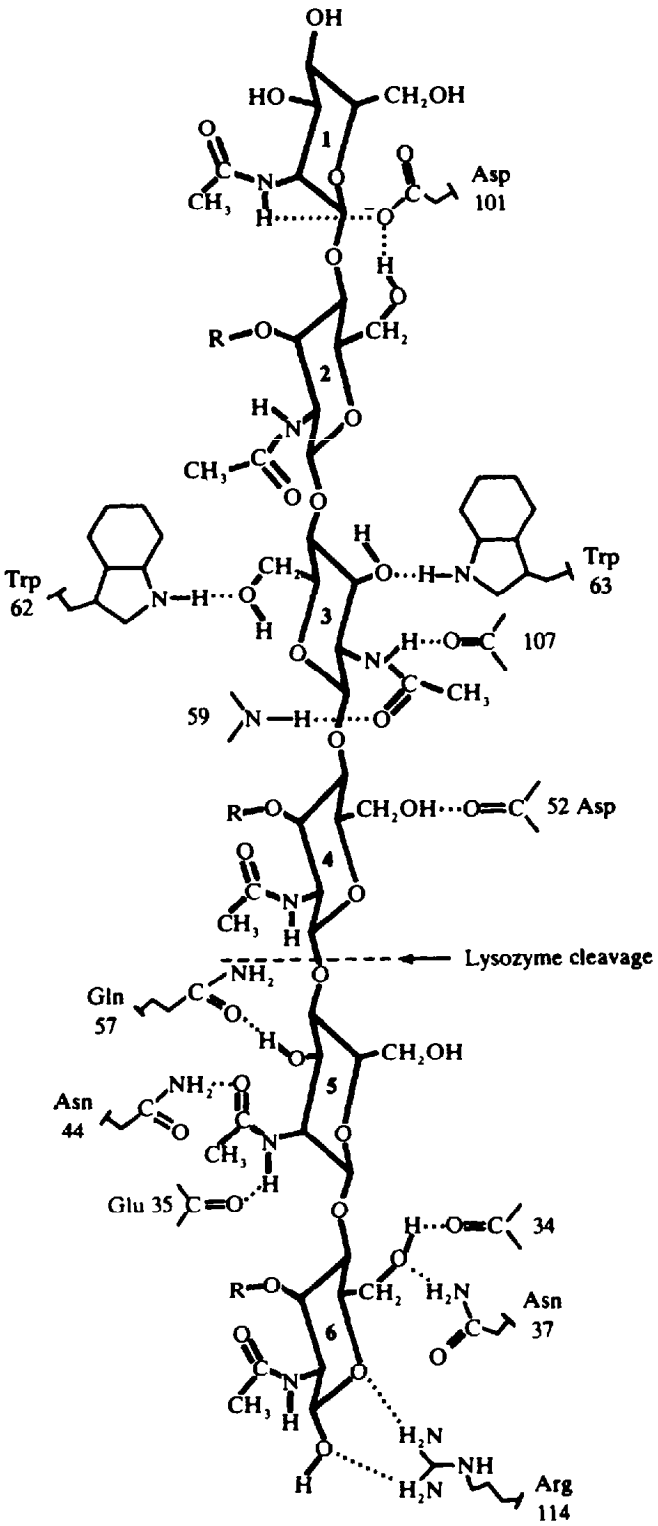


Fig. 8-2 Binding of the substrate $(NAG\text{-}NAM)_3$ to the active site of lysozyme. The substrate is drawn with bold bonds, the enzyme groups with light-face bonds. H bonds are indicated by dotted lines.

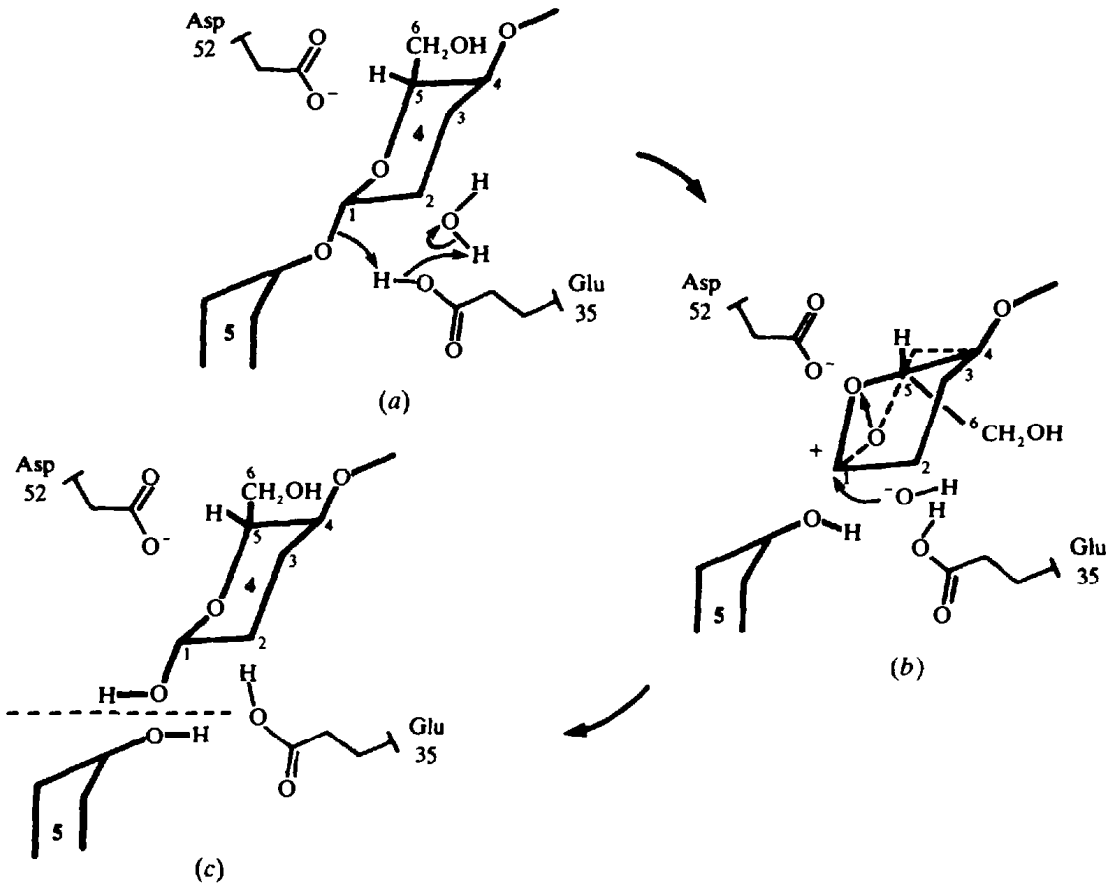


Fig. 8-3 Probable mechanism of bond cleavage by lysozyme. The thick solid line structure is the substrate, and the lightface groups are on the enzyme. The small arrows indicate displacement of electron pairs during the reaction.

(*b*) The carboxyl of Glu 35 donates a proton, cleaving the C-1—O bond and releasing the disaccharide [5]-[6]. The resulting carbocation, C-1 of ring [4], is stabilized by the negatively charged Asp 52. The carbocation then reacts with the OH^- from the solvent water to release the tetrasaccharide [1]-[2]-[3]-[4]. Glu 35 is then reprotonated in readiness for the *next* round of reactions. The glutamic acid acts as a proton donor for the reaction, which is thus classified as general acid catalysis.

(*c*) Since proton donation is *central* to the catalytic process, this is an example of general acid-base catalysis; specifically, it is acid catalysis and it is noncovalent.

(*d*) In Fig. 8-3, the numerous dotted lines drawn between O's and NH's of amino acid residues and O's and NH's of the oligosaccharide indicate hydrogen bonding; however, van der Waals, noncovalent bonding also occurs.

(*e*) The binding of the substrate distorts the previous chair conformation of residue [4]; this *reduces* the tendency for binding; i.e., ΔG for binding will be elevated. However, this energy of distortion (strain) contributes to the total activation energy required for subsequent bond cleavage via formation of the carbonium ion.

8.11. *Carboxypeptidase A* (EC 3.4.17.1) is a pancreatic digestive enzyme consisting of a single polypeptide chain of 307 amino acids with a total M_r of 36,000. It catalyzes the cleavage of amino acid residues from C termini of polypeptides. Importantly, for its mechanisms of action,

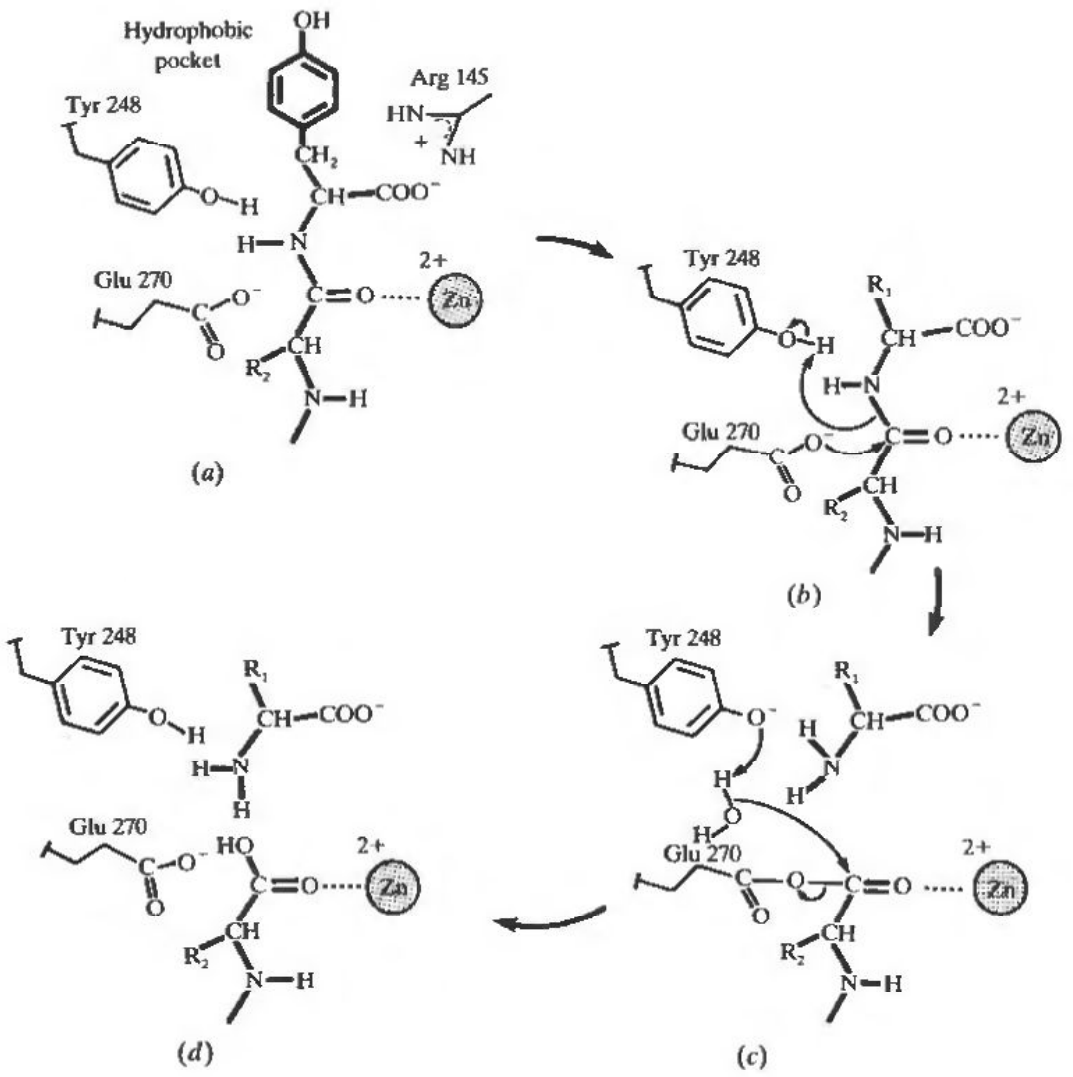


Fig. 8-4 The mechanism of covalent catalysis of the hydrolysis of a C-terminal amino acid residue from a peptide by carboxypeptidase A. The reaction is (*a*)→(*d*), and the bold line structure is the peptide substrate. The C-terminal tyrosine side chain of the substrate shown in (*a*) is denoted by R_1 in (*b*), (*c*), and (*d*).

it contains one Zn^{2+} in its active site. The amino acid side chains that form its active site and the catalytic sequence are shown in Fig. 8-4.

(*a*) To what general class of enzyme does carboxypeptidase belong?

(*b*) What basic role does the Zn^{2+} play in the catalytic mechanism? What general type of catalysis occurs?

(*c*) Describe in words the sequence of events that is depicted in Fig. 8-4.

SOLUTION

(*a*) Carboxypeptidase is a hydrolase; i.e., it catalyzes the hydrolytic cleavage of a (peptide) bond. The second integer in its EC number sequence indicates that it cleaves a C—N bond. It is an *exopeptidase*; i.e., it hydrolyzes amino acid residues from the carboxyl termini of peptides. There exist also *amino*peptidases that catalyze the hydrolytic removal of N-terminal amino acid residues. *Endo*peptidases are those hydrolases that hydrolyze peptide bonds, not at the C or N termini, but *within* the chain; examples are pepsin in the stomach and the pancreatic peptidases, such as chymotrypsin and trypsin.

(*b*) The Zn^{2+} acts as an electrophile that further polarizes the carbonyl oxygen [Fig. 8-4(*c*)] before the formation of an ester linkage with the γ carboxyl of Glu 270 of the enzyme. This linkage is covalent, so this reaction is an example of covalent catalysis.

(*c*) (1) The peptide substrate binds to the active site, which contains Arg 145, the Zn^{2+} ion, and the so-called *hydrophobic pocket*, which contains aromatic and aliphatic amino acid side chains. (2) Nucleophilic attack by Glu 270 on the peptide bond is accompanied by the uptake of an H^+ from Tyr 248. (3) This results in cleavage of the peptide bond and diffusion away of the C-terminal free amino acid. (4) The covalently bound polypeptide is then released to regenerate free enzyme by nucleophilic attack of an H_2O molecule on the anhydride bond of Glu 270; this is followed by reprotonation of Tyr 248.

RATE ENHANCEMENT AND ACTIVATION ENERGY

8.12. Calculate the rate enhancement that would be achieved if the activation energy of the transition-state complex of an enzyme with its substrate were halved.

SOLUTION

From Eq. (*8.2*) k is proportional to $e^{-\Delta G^{\ddagger}/RT}$. Thus,

$$k_{NEW} = k_{OLD} \exp(-\Delta G^{\ddagger}/2RT)$$

It is thus clear that the rate enhancement is dependent on both the original $\Delta G^{\ddagger}$ *and* the temperature. So, unless these values are given, the enhancement cannot be calculated.

8.13. If $\Delta G^{\ddagger}$ and T in the previous example were -1 kJ mol^{-1} and 300 K, respectively, what would be the rate enhancement factor?

SOLUTION

$$k_{NEW} = k_{OLD} \exp\,[1{,}000/(2 \times 8.314 \times 300)] = k_{OLD}\ 1.22$$

Thus, the rate enhancement would be 22 percent. Clearly, enzymes achieve far more dramatic rate increases than this. This suggests that $-\Delta G^{\ddagger}$ values are much larger in the first place (namely, in free solution) and they are more dramatically reduced (than the twofold reduction discussed in this problem) in enzymes.

Supplementary Problems

8.14. Draw the structures of the substrates and products of the following reactions and give the general EC classification of the enzyme involved:

(*a*) Glucose 6-phosphate + $H_2O \xrightarrow[\text{Glucose 6-phosphatase}]{}$ . . . . . .

(*b*) Lactate + $NAD^+ \underset{\text{Lactate dehydrogenase}}{\rightleftharpoons}$ NADH + H^+ + ...

(*c*) Argininosuccinate $\underset{\text{Argininosuccinase}}{\rightleftharpoons}$ Fumarate + ...

(*d*) Fumarate + $H_2O \xrightarrow[\text{Fumarase}]{}$ . . . . . .

8.15. Alcohol dehydrogenase catalyzes the oxidation of a variety of alcohols to their corresponding aldehydes. For a given amount of enzyme and a given substrate concentration (mol L^{-1}), the rates of reaction with the following substrates differ: methanol, ethanol, propanol, butanol, cyclohexanol, phenol. (*a*) Arrange the substrates in the order of decreasing reaction rate. (*b*) Give reasons for your speculations.

8.16. Glyceraldehyde-3-phosphate dehydrogenase has an essential cysteine residue in its active site. The enzyme forms a transient acyl compound with its substrate, glyceraldehyde 3-phosphate. (*a*) What is the general chemical name of the compound? (*b*) Draw its likely structure.

8.17. What is the basic difference between the reaction catalyzed by (*a*) a *mutase* enzyme and that catalyzed by an *isomerase*; (*b*) an *oxidase*, an *oxygenase*, and the reverse reaction of a *reductase*? Give examples.

8.18. Given that the spontaneous hydration of CO_2 is reasonably fast (Example 8.4), what might be a physiological rationalization for the need for the enzyme carbonic anhydrase?

8.19. What is a *suicide* substrate of an enzyme?

8.20. If the enzyme concentration in a reaction mixture at equilibrium is comparable with that of the reactants, is the ratio of product concentrations to substrate concentrations the same as if no enzyme were present?

8.21. Why are most enzymes so large relative to their substrates?

8.22. Give some examples of enzymes that are smaller than their substrates.

Chapter 9

Enzyme Kinetics

9.1 INTRODUCTION AND DEFINITIONS

Enzyme kinetics is concerned with measuring the rates of enzymatic reactions, and with factors that affect the rates.

Question: What determines the rates of enzymatic reactions, and why is it important?

The important factors influencing the rates of enzymatic reactions are the concentrations of substrates and enzyme, as well as factors such as pH, temperature, and the presence of cofactors and metal ions. A study of these factors is important. In a pragmatic sense, there may be occasions when one needs to optimize the rate of a particular reaction. A study of the way the rate depends on experimental variables may also allow one to discriminate between possible models that attempt to predict how the enzyme may function, i.e., its mechanism.

In addition to the experimental aspects of enzyme kinetics, design of experiments, and methods for determining the progress of enzymatic reactions, an important aspect is the *interpretation* of the data. This usually depends on writing mathematical expressions for model reaction schemes, which predict how the rate depends on reaction variables. These equations are then tested for consistency with experimental data, which may allow the rejection of models that do not satisfactorily predict the measured behavior.

The basic principles and definitions used frequently in enzyme kinetic analysis are discussed below.

Principle of Mass Action

For a single, irreversible step in a chemical reaction, i.e., an *elementary* chemical process, the rate of the reaction is proportional to the concentrations of the reactants involved in the process. The constant of proportionality is called the *rate constant*, or the *unitary rate constant* to highlight the fact that it applies to an elementary process. A subtlety that may be introduced into rate expressions is to use *chemical activities* (see Chap. 10) and not simply concentrations, but activity coefficients in biological systems are generally taken to be near 1.

EXAMPLE 9.1

Application of the principle of mass action to the reaction scheme

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P + Q$$

with forward and reverse rate constants k_1 and k_{-1}, leads to the following expressions for the forward and reverse reaction rates:

$$\begin{aligned} \text{forward rate} &= k_1[A][B] \\ \text{reverse rate} &= k_{-1}[P][Q] \end{aligned} \tag{9.1}$$

where the square brackets denote concentration in mol L^{-1}. At *chemical equilibrium*, the forward and reverse rates are equal, so there is no *net* production of any of the reactants with time. Thus,

$$\frac{k_1}{k_{-1}} = \frac{[\text{P}]_e[\text{Q}]_e}{[\text{A}]_e[\text{B}]_e} = K_e \tag{9.2}$$

where K_e is termed the *equilibrium constant* and the subscript e denotes the equilibrium value of the concentrations.

Reaction rates are simply concentration changes of a species per unit of time and therefore can be written mathematically as *derivatives*. Note, however, that the mathematical expression for the overall rate of change of, say, [A] must include forward and reverse *fluxes* (Greek: "to flow"); for example,

$$\frac{d[\text{A}]}{dt} = -k_1[\text{A}][\text{B}] + k_{-1}[\text{P}][\text{Q}] \tag{9.3}$$

Molecularity

Molecularity refers to the *number of molecules* involved in an elementary reaction. Usually, only two molecules collide in one instant to give product(s) (molecularity = 2) or a single molecule undergoes *fission* (also called *scission*; molecularity = 1). Example 9.1 is of a reaction in which the forward and reverse processes have a molecularity of two.

Order of a Reaction

This is the *sum of the powers* to which the concentration (or chemical activity) terms are raised in a rate expression.

EXAMPLE 9.2

In the first-order reaction

$$\text{A} \xrightarrow{k} \text{P}$$

the expression for the *rate of change* of [A] is

$$\frac{d[\text{A}]}{dt} = -k[\text{A}] \tag{9.4}$$

Since the left-hand side of the expression has the units-of-reaction rate ($\text{mol L}^{-1}\text{s}^{-1}$), then these units must also apply to the right hand side (balance of dimensions). Therefore, the units of $k[\text{A}]$ must be $\text{mol L}^{-1}\text{s}^{-1}$, implying that k has units of s^{-1}. Thus, simple *dimensional analysis* leads directly to the general expression for the units of a particular constant in a particular reaction scheme.

9.2 DEPENDENCE OF ENZYME REACTION RATE ON SUBSTRATE CONCENTRATION

Experimentally, the effect of substrate concentration on enzyme reaction rate is studied by recording the progress of an enzyme-catalyzed reaction, using a fixed concentration of enzyme and a series of different substrate concentrations. The *initial velocity*, v_0, is measured as the slope of the tangent of the progress curve at time $t = 0$. The initial velocity is used because enzyme degradation during the reaction or inhibition by reaction products may occur, thus yielding results that may be difficult to interpret.

When $[\text{S}]_0 \gg$ the enzyme concentration, v_0 is usually directly proportional to the enzyme concentration in the reaction mixture, and for most enzymes v_0 is a rectangular hyperbolic function

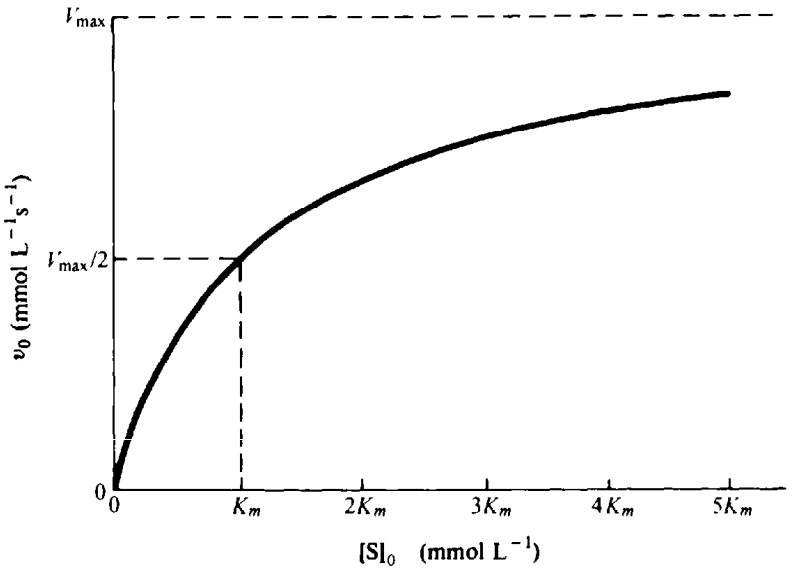


Fig. 9-1 The hyperbolic relationship between initial velocity (v_0) and initial substrate concentration ($[S]_0$) of an enzyme-catalyzed reaction.

of $[S]_0$ (Fig. 9-1). If there are other (co-) substrates, then these are usually held constant during the series of experiments in which $[S]_0$ is varied.

The equation describing the rectangular hyperbola that usually represents enzyme-reaction data (e.g., Fig. 9-1) is called the *Michaelis-Menten equation*:

$$v_0 = -\left(\frac{d[S]}{dt}\right)_{t=0} = \frac{V_{max}[S]_0}{K_m + [S]_0} \tag{9.5}$$

The equation has the property that when $[S]_0$ is very large, $v_0 = V_{max}$ (the so-called *maximal velocity*); also when $v_0 = V_{max}/2$, the value of $[S]_0$ is K_m, the so-called *Michaelis constant*.

9.3 GRAPHICAL EVALUATION OF K_m AND V_{max}

Eq. (*9.5*) can be rearranged into several new forms that yield straight lines when one *new* variable is plotted against the other. The advantages of this mathematical manipulation are that (1) V_{max} and K_m can be determined readily by fitting a straight line to the transformed data; (2) departures of the data from a straight line are more easily detected than is nonconformity to a hyperbola (these departures may indicate an inappropriateness of the simple enzyme model); (3) the effects of inhibitors on the reaction can be analyzed more easily. It is also easier and more accurate to obtain an extrapolated intercept from a linear plot than to estimate by eye an asymptote to a rectangular hyperbola.

The most commonly used transformation of the Michaelis-Menten equation is the Lineweaver-Burk "double reciprocal" equation.

This equation was first introduced in 1935. By taking reciprocals of both sides of Eq. (*9.5*) we obtain

$$\frac{1}{v_0} = \frac{K_m}{V_{max}}\frac{1}{[S]_0} + \frac{1}{V_{max}} \tag{9.6}$$

A plot of data pairs $(1/[S]_{0,i}, 1/v_{0,i})$, for $i = 1, \ldots, n$, where n is the number of data pairs, gives a straight line with ordinate and abscissa intercepts $1/V_{max}$ and $-1/K_m$, respectively [Fig. 9-2].

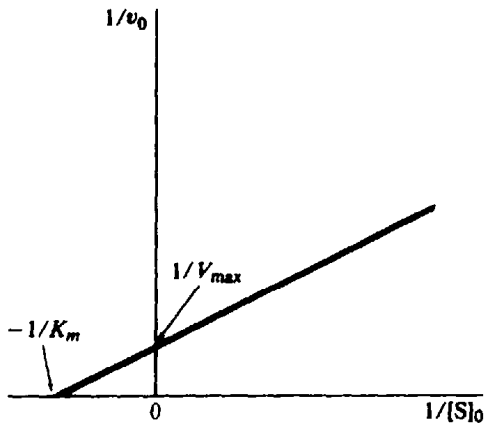


Fig. 9-2 Lineweaver-Burk graphical procedure for determining the two steady-state kinetic parameters in the Michaelis-Menten equation.

Question: Are graphical procedures still used to obtain the best estimates of steady-state parameters?

With the general availability of computers, graphical methods have to a large extent been replaced by the procedure of *nonlinear least squares regression* of the kinetic equation *directly* onto the *untransformed* data. Nevertheless, graphical procedures are still valuable for obtaining a quick guide to the parameter values or for obtaining the initial estimates that are required by the numerical methods.

9.4 ENZYME INHIBITION—DEFINITIONS

Often, rates of enzymatic reactions are affected by substances that are not reactants; when there is a reduction in rate caused by a compound, then the compound is said to be an *inhibitor*. Increased reaction rate by an *activator* is the opposite of this effect. In dealing with inhibitors it is important to distinguish between the *effects* that are observed experimentally and the *mechanisms* (or models) proposed to explain them.

There are three basic types of inhibition. They are defined in terms of the *degree of inhibition* i, which itself is defined as

$$i = \frac{v_0 - v_i}{v_0} \tag{9.7}$$

where v_0 and v_i are the uninhibited and inhibited initial reaction rates, respectively.

1. *Pure noncompetitive inhibition* is said to exist if i is unaffected by the concentration of substrate.
2. *Competitive inhibition* exists if i decreases as the substrate concentration is increased.
3. *Anti- or uncompetitive inhibition* exists if i increases as the substrate concentration is increased.

In addition to the above, there is *mixed inhibition*, which exists if i increases or decreases as substrate concentration increases, but *not to the same extent* as for the pure competitive or anticompetitive cases,

respectively. In fact, it can be shown that *mechanistically*, noncompetitive inhibition is a special case of mixed inhibition; however, *operationally*, as defined here, mixed inhibition is a combination of two of the three basic types.

9.5 ENZYME INHIBITION—EQUATIONS

The mathematical expressions relating reaction rate and inhibitor concentration are often rather complicated, but there are four simple equations that are extensions of the Michaelis-Menten formula. These merit special consideration because the kinetics of many enzymes can be satisfactorily described by them. In the equations in Table 9.1, [I] denotes the inhibitor concentration and K_I and K_I' are inhibition constants, the units of which are those of a dissociation equilibrium constant (mmol L^{-1}). Mechanisms that are consistent with these equations are described in Sect. 9.10.

9.6 MECHANISTIC BASIS OF THE MICHAELIS-MENTEN EQUATION

Equilibrium Analysis

To explain their results on the conversion of sucrose to glucose and fructose by the enzyme *invertase*, Michaelis and Menten proposed in 1913 the following scheme of reactions:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

It was assumed that k_{-1} is large, compared with k_2; thus, the first part of the reaction could be described by the equilibrium constant for the dissociation of the enzyme-substrate complex: $K_s = [E][S]/[ES]$. The concentrations of S and E at any time are derived from the known *initial conditions*:

$$[S]_0 = [S] + [ES] + [P] \tag{9.8}$$

$$[E]_0 = [E] + [ES] \tag{9.9}$$

Table 9.1. Rate Equations for the Four Types of Enzyme Inhibition

Pure Noncompetitive	Pure Competitive
$v_i = \dfrac{V_{max}[S]_0}{(K_m + [S]_0)\left(1 + \dfrac{[I]}{K_I}\right)}$ $i = \dfrac{[I]}{K_I + [I]}$	$v_i = \dfrac{V_{max}[S]_0}{K_m\left(1 + \dfrac{[I]}{K_I}\right) + [S]_0}$ $i = \dfrac{K_m \dfrac{[I]}{K_I}}{K_m\left(1 + \dfrac{[I]}{K_I}\right) + [S]_0}$
Anticompetitive	**Mixed**
$v_i = \dfrac{V_{max}[S]_0}{K_m + \left(1 + \dfrac{[I]}{K_I}\right)[S]_0}$ $i = \dfrac{[S]_0 \dfrac{[I]}{K_I}}{K_m + \left(1 + \dfrac{[I]}{K_I}\right)[S]_0}$	$v_i = \dfrac{V_{max}[S]_0}{K_m\left(1 + \dfrac{[I]}{K_I}\right) + \left(1 + \dfrac{[I]}{K_I'}\right)[S]_0}$ $i = \dfrac{[I]\left\{\dfrac{K_m}{K_I} + \dfrac{[S]_0}{K_I'}\right\}}{K_m\left(1 + \dfrac{[I]}{K_I}\right) + \left(1 + \dfrac{[I]}{K_I'}\right)[S]_0}$

Since experimentally, $[S]_0 \gg [E]_0$, then $[S] \approx [S]_0$ early in the reaction, and by using the expression for K_s and Eq. (9.9), we obtain:

$$K_s = \frac{([E]_0 - [ES])[S]_0}{[ES]} \tag{9.10}$$

This equation can be rearranged to give

$$[ES] = \frac{[E]_0[S]_0}{K_s + [S]_0} \tag{9.11}$$

The second step of the reaction is a simple first-order one, and thus

$$v_0 = k_2[ES] \tag{9.12}$$

Hence the overall rate of decline of substrate is described by

$$v_0 = \frac{k_2[E]_0[S]_0}{K_s + [S]_0} \tag{9.13}$$

which is exactly the form of Eq. (9.5), if $V_{max} = k_2[E]_0$ and $K_m = K_s$.

Steady-State Analysis

Briggs and Haldane in 1925 examined the earlier Michaelis-Menten analysis and made an important development. Instead of assuming that the first stage of the reaction was at equilibrium, they merely assumed, for all intents and purposes, that the concentration of the enzyme-substrate complex scarcely changed with time; i.e., it was in a *steady state*. Written mathematically, this amounts to

$$\frac{d[ES]}{dt} = 0 \tag{9.14}$$

Now, the flux equation for [ES] is

$$\frac{d[ES]}{dt} = k_1[E][S] - (k_{-1} + k_2)[ES] \tag{9.15}$$

and by using Eqs. (9.8) and (9.9) with $[S] \approx [S_0]$, then

$$0 = k_1[E]_0[S]_0 - (k_1[S]_0 + k_{-1} + k_2)[ES] \tag{9.16}$$

By rearranging this equation and using the fact that $v_0 = k_2[ES]$, we obtain

$$v_0 = \frac{k_2[E]_0[S]_0}{\frac{(k_{-1} + k_2)}{k_1} + [S]_0} \tag{9.17}$$

Again, this equation has the same form as Eq. (9.5), provided K_m is *identified with* $(k_{-1} + k_2)/k_1$ and V_{max} is identified with $k_2[E]_0$.

Several features of Eq. (9.17) are worth noting:

1. Because k_2 describes the number of molecules of substrate converted to product per second per molecule of enzyme, it is called the *turnover number* of the enzyme. Generally, in more complex enzyme mechanisms, the expression for V_{max} is complicated by k_2 being replaced by an expression that is a ratio of sums of products of unitary rate constants; this grouped expression is then called k_{cat}.
2. If an enzyme is not pure, it may not be possible to determine accurately the concentration of the active from, $[E]_0$. Nevertheless, V_{max} can still be obtained by steady-state kinetic

analysis. So, to standardize experimental results, we refer to *one enzyme unit* (or *katal*) as the amount of enzyme solution required to transform 1 μmol of substrate into product(s) in 1 min, under standard conditions of pH, ionic strength, and temperature.

3. When $[S]_0$ is very large compared with K_m, virtually all of E is in the form of ES, so the enzyme is said to be *saturated*, i.e., it is then operating at its maximum velocity.

9.7 DERIVATION OF COMPLICATED STEADY-STATE EQUATIONS

In principle, the steady-state rate expression for any enzyme with any number of reactants can be derived using the methods of the previous section. In practice, the procedure is very laborious, so use is made of an *algorithmic method*, introduced by King and Altman in 1956; it is *not* applicable to (1) nonenzymatic reactions (each reactant concentration must be $\gg[E]_0$), (2) mixtures of enzymes, or (3) reactions with nonenzymatic steps. However, these are not severe restrictions. It is applied as follows:

1. Draw the reaction scheme (the *master pattern*) with the required reaction arrows interconnecting all relevant enzyme species (free and complexed forms).
2. Annotate all reaction arrows with the corresponding unitary rate constant. For forward reactions where a substrate is involved, place its letter of designation next to the rate constant in the scheme; do the same for reverse reactions involving a product.
3. For the n enzyme forms (one of free enzyme and the rest complexes or isomeric forms of the enzyme), draw *reaction patterns* that have $n-1$ arrows and yield a continuous path or paths that lead to each enzyme form. In addition, *no closed loops* of steps are allowed in the pattern.
4. The expression for $[ES_i]\cdot[E]_T$, where $[ES_i]$ is any enzyme form and $[E]_T$ is the *total* enzyme concentration, is given by the summed products of concentrations and rate constants from each pattern.
5. The expression for the overall rate of *product formation* alone is given by multiplying the concentration of the relevant enzyme-substrate complex (say, $[ES_k]$) and the first-order rate constant of its breakdown (say k_j):

$$v = \frac{k_j[E]_T[ES_k]}{\begin{array}{c}\text{(sum of all expressions for [E] and the}\\ \text{different } [ES_i],\ \text{where } i = 1,\ldots,n-1)\end{array}} \tag{9.18}$$

In applying the algorithmic method, note that:

(*a*) In all mechanisms not involving alternative reaction sequences, the numerator of the rate expression will have only two terms: one, the product of all *forward* rate constants and substrate concentrations [as in Eq. (*9.18*)] and the other, the corresponding product for the *reverse* reaction.

(*b*) With complicated mechanisms, it is easy to overlook some patterns at step 3, above. A useful formula that yields the total number of $(n-1)$-line patterns for m steps, where n is the number of enzyme forms, is

$$\text{Total } (n-1)\text{-line patterns} = \binom{m}{n-1} = \frac{m!}{(n-1)!(m-n+1)!} \tag{9.19}$$

(*c*) Eq. (*9.19*) predicts the number of $(n-1)$-line patterns, of which some may have closed loops; the closed-loop patterns must be eliminated. The total number of $(n-1)$-line patterns with r-sided closed loops, Z, is

$$Z = \begin{pmatrix} m-r \\ n-1-r \end{pmatrix} = \frac{(m-r)!}{(n-1-r)!(m-n+1)!} \tag{9.20}$$

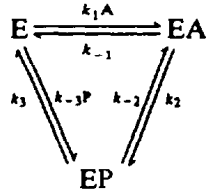
Hence, Z is determined for a range from 1 to $n-1$, and the sum of all these Z values is the total number of excluded patterns.

EXAMPLE 9.3

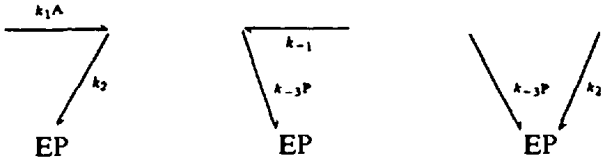
By using the King-Altman procedure, derive the steady-state rate equation for the following enzyme mechanism:

$$A + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EA \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} E + P$$

Steps (1) and (2): The master pattern is:



Step (3): The reaction patterns for EP are:



Step (4): Therefore,

$$[EP] = k_1k_2[A] + k_{-1}k_{-3}[P] + k_2k_{-3}[P] \tag{9.21}$$

Similarly, the expressions for the other two enzyme species are

$$[EA] = k_1k_{-2}[A] + k_{-2}k_{-3}[P] + k_1k_3[A]$$
$$[E] = k_{-1}k_3 + k_2k_3 + k_{-1}k_{-2}$$

Since

$$v = \frac{d[P]}{dt} = k_3[EP] - k_{-3}[E][P] \tag{9.22}$$

and

$$[E]_T = [E] + [EA] + [EP] \tag{9.23}$$

then

$$v = \frac{[E]_T\{k_3(k_1k_2[A] + k_{-1}k_{-3}[P] + k_2k_{-3}[P]) - k_{-3}[P](k_{-1}k_3 + k_2k_3 + k_{-1}k_{-2})\}}{k_1k_2[A] + k_{-1}k_{-3}[P] + k_2k_{-3}[P] + k_1k_{-2}[A] + k_{-2}k_{-3}[P] + k_1k_3[A] + k_{-1}k_3 + k_2k_3 + k_{-1}k_{-2}} \tag{9.24}$$

which is simplified to

$$v = \frac{[E]_T(k_1k_2k_3[A] - k_{-1}k_{-2}k_{-3}[P])}{(k_{-1}k_3 + k_2k_3 + k_{-1}k_{-2}) + k_1(k_2 + k_3 + k_{-2})[A] + k_{-3}(k_{-1} + k_2 + k_{-2})[P]} \tag{9.25}$$

If [P] = 0, then the expression of Eq. (*9.25*) is of the form

$$v = \frac{[E]_T(\text{num1})[A]}{\text{coef} + \text{coefA}[A]} \quad \text{or} \quad v = \frac{[E]_T\left(\dfrac{\text{num1}}{\text{coefA}}\right)[A]}{\left(\dfrac{\text{coef}}{\text{coefA}}\right) + [A]} \tag{9.26}$$

where the numerator coefficient is designated num1 and the two denominator coefficients are designated coef and coefA, respectively. When the numerator and denominator are divided by coefA, the form of Eq. (*9.26*) is identical to the Michaelis-Menten equation, Eq. (*9.5*). Thus, the analysis illustrates a very important rule of steady-state enzyme kinetics: The introduction to a mechanism of steps that involve only isomerization between enzyme complexes (EP in this case) does not change the *form* (see next section) of the steady-state rate equation.

9.8 MULTIREACTANT ENZYMES

General

The most common enzymatic reactions are those with two or more substrates and as many products. But many of the simpler single-substrate schemes are valuable for the development of kinetic ideas concerning effects of pH, temperature, etc., on enzyme reaction rates. Although the mechanisms of multisubstrate reactions are complicated, their kinetics can often be described by an equation of the form:

$$v = \frac{V_{\max}^{\text{app}}[A]}{K_m^{\text{app}} + [A]} \tag{9.27}$$

This is so if the concentrations of all substrates other than A are held constant during the experiments; the values of $V_{\max}^{\text{app}}$ and K_m^{app} are functions of the concentrations of the other reactants (Example 9.6).

Nomenclature

In 1963 W. W. Cleland published a classification of enzyme-catalyzed reactions based on the number of substrates and products in the reaction. This classification is as follows:

1. The *reactancy* is the number of kinetically significant substrates or products and is designated by the syllables *Uni*, *Bi*, *Ter*, *Quad*.

EXAMPLE 9.4

The associated Cleland designations for the reactions below are

$A \rightleftarrows P$	Uni Uni
$A \rightleftarrows P + Q$	Uni Bi
$A + B \rightleftarrows P + Q$	Bi Bi
$A + B + C \rightleftarrows P + Q + R + S$	Ter Quad

2. If all substrates add to the enzyme *before* any products are released, the mechanism is defined as *sequential*. If substrates add in an *obligatory* order, the mechanism is called *sequential ordered*. If there is no obligatory order of addition of substrates or release of products, it is called *sequential random*. When one or more products are released before all the substrates have been added, the enzyme will exist in two or more stable forms between which it oscillates during the reaction; this type of mechanism is therefore called *Ping Pong*. If *isomerization* of stable, as distinct from transitory (e.g., EAB EPQ in Example 9.5), enzyme forms occurs, the term *Iso* is added to the designation of the mechanism.

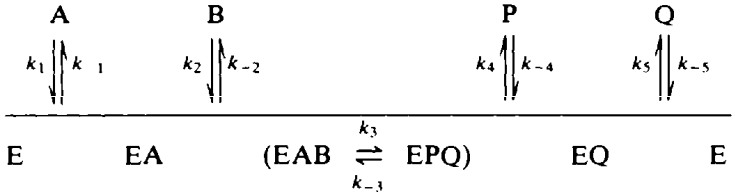
EXAMPLE 9.5

The ordered Bi Bi mechanism can be written as

$$E + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EA \qquad EA + B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EAB \qquad EAB \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} EPQ$$

$$EPQ \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} EQ + P \qquad EQ \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}} E + Q$$

or simply, using Cleland's diagrammatical convention,* as



In Cleland's convention the letters A, B, C, D denote substrates and P, Q, R, S denote products, in the order in which they add and leave the enzyme, respectively.

EXAMPLE 9.6

By using the King-Altman procedure, the rate expression for the ordered Bi Bi mechanism in the previous example can be shown to be:

$$\begin{aligned} v = {} & (k_1k_2k_3k_4k_5[A][B] - k_{-1}k_{-2}k_{-3}k_{-4}k_{-5}[P][Q]) \div \{k_5k_{-1}(k_3k_4 + k_4k_{-2} + k_{-2}k_{-3}) \\ & + [A]k_1k_5(k_3k_4 + k_4k_{-2} + k_{-2}k_{-3}) + [B]k_2k_3k_4k_5 + [P]k_{-1}k_{-2}k_{-3}k_{-4} \\ & + [Q]k_{-1}k_{-5}(k_3k_4 + k_4k_{-2} + k_{-2}k_{-3}) + [A][B]k_1k_2(k_3k_4 + k_3k_5 + k_4k_5 + k_5k_{-3}) \\ & + [P][Q]k_{-4}k_{-5}(k_3k_{-1} + k_{-1}k_{-2} + k_{-1}k_{-3} + k_{-2}k_{-3}) + [A][P]k_1k_{-2}k_{-3}k_{-4} \\ & + [B][Q]k_2k_3k_4k_{-5} + [A][B][P]k_1k_2k_{-4}(k_3 + k_{-3}) + [B][P][Q]k_2k_{-4}k_{-5}(k_3 + k_{-3})\} \end{aligned} \tag{9.28}$$

The expression is simplified dramatically if we assume that at the start of the reaction [P] and [Q] = 0. Furthermore, if [B] is *saturating*, we can divide both the numerator and denominator by [B] and take the limit as $[B] \to \infty$, and the equation reduces to

$$v_0 = \frac{\text{num1}[A]_0}{\text{coefB} + \text{coefAB}[A]_0} \tag{9.29}$$

where the terms num1 and coefAB and coefB relate to Eq. (*9.28*). Therefore, the expressions for V_{max} and K_m, respectively, are given by

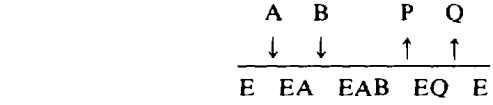
$$V_{max} = \frac{\text{num1}}{\text{coefAB}} \tag{9.30}$$

$$K_m = \frac{\text{coefB}}{\text{coefAB}} \tag{9.31}$$

Expressions for the parameters relating to B may be similarly derived.

---

*In its original form, Cleland's notation had only *single* arrows; thus this mechanism would be written as



9.9 pH EFFECTS ON ENZYME REACTION RATES

General

Question: What is so useful about studying the effect of pH on the rates of enzymatic reactions?

Pragmatically, one may simply want to find the pH at which the enzyme is most efficient at catalyzing the reaction: the *optimum pH*. In addition, a thorough and careful study of the effects of pH on the kinetics of an enzymatic reaction may lead to an understanding of which functional groups of the enzyme or substrate may be involved in binding or catalysis.

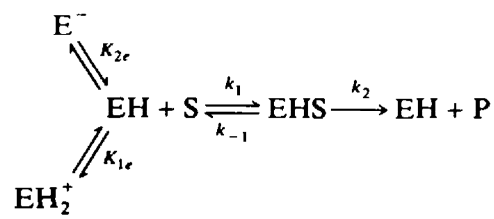
To describe completely the effects of pH changes on enzyme catalysis is an almost impossible task. Many of the amino acid side chains in an enzyme are ionizable, but in environments with polarities different from that of the free solution, their pK_a's (Chap. 3) will probably be significantly altered. However, experimentally, it is a simple matter to determine values of steady-state kinetic parameters (K_m, V_{max}) of an enzyme for various pH conditions.

The possible effects of pH are to change the ionization state of (1) groups involved in catalysis, (2) groups involved in the binding of substrate, (3) groups involved in binding at sites other than the active site, defined as *allosteric effector sites*, and (4) groups on the substrates. These altered charge states will affect the affinity of the enzyme for its substrates and the rate of catalysis (Chap. 8).

A Simple Model of pH Effects

Ionization of Free Enzyme

The simplest Michaelis-Menten-type scheme for showing analytically the effect of enzyme ionization on enzyme kinetic parameters is



where $$K_{2e} = \frac{[E^-][H^+]}{[EH]} \qquad K_{1e} = \frac{[EH][H^+]}{[EH_2^+]} \qquad K_s = \frac{[EH][S]}{[EHS]} \tag{9.32}$$

The rate expression corresponding to this scheme is derived most readily by using the equilibrium analysis of Michaelis and Menten. The *conservation of mass equation* for the enzyme is

$$[E]_0 = [E^-] + [EH] + [EHS] + [EH_2^+] \tag{9.33}$$

By using the relationship in Eq. (*9.33*), we obtain

$$[EHS] = \frac{[E]_0}{1 + \dfrac{K_s}{[S_0]}\left(1 + \dfrac{K_{2e}}{[H^+]} + \dfrac{[H^+]}{K_{1e}}\right)} \tag{9.34}$$

Since $v_0 = k_2[EHS]$, then

$$v_0 = \frac{k_2[E]_0[S]_0}{K_s\left(1 + \dfrac{K_{2e}}{[H^+]} + \dfrac{[H^+]}{K_{1e}}\right) + [S]_0} \tag{9.35}$$

Eq. (*9.35*) corresponds to the Michaelis-Menten equation with the same expression for V_{max}, $k_2[E]_0$, but

$$K_m = K_s\left(1 + \frac{K_{2e}}{[H^+]} + \frac{[H^+]}{K_{1e}}\right) \tag{9.36}$$

The following facts about these equations should be noted:

(*a*) For fixed values of $[S]_0$ and $[E]_0$, Eq. (*9.35*) describes a bell-shaped curve with a maximum at a pH between pK_{1e} and pK_{2e}. In other words, a plot of $-\log$(apparent K_m) $\equiv pK_m$ versus pH has the form shown in Fig. 9-3, with a maximum value of pK_m at $pH = -1/2\log(K_{1e} \cdot K_{2e})$; i.e., the optimum pH for the reaction is $(pK_{1e} + pK_{2e})/2$.

(*b*) Since at any pH the term in parentheses in Eq. (*9.36*) is always >1, then $K_m > K_s$.

(*c*) The concentration of H^+ (i.e., pH) has no effect on V_{max}.

(*d*) If $[H^+] \gg K_{2e}$, that is, when the pH is low (i.e., $[H^+]$ is high), then Eq. (*9.35*) becomes

$$v_0 = \frac{V_{max}[S]_0}{K_s\left(1 + \dfrac{[H^+]}{K_{1e}}\right) + [S]_0} \tag{9.37}$$

which is the same form as the competitive inhibition equation (Table 9.1) with H^+ as the *inhibitor*.

(*e*) At low pH, $K_m \simeq K_s[H^+]/K_{1e}$; hence, $pK_m = pK_s + pH - pK_{1e}$, and a plot of pK_m versus pH is a straight line with a slope of 1 (Fig. 9-3). Similarly, at high pH, where $K_{1e} \gg K_{2e} \gg [H^+]$, $K_m = K_s K_{2e}/[H^+]$ and $pK_m = pK_s + pK_{2e} - pH$, and a plot of pK_m versus

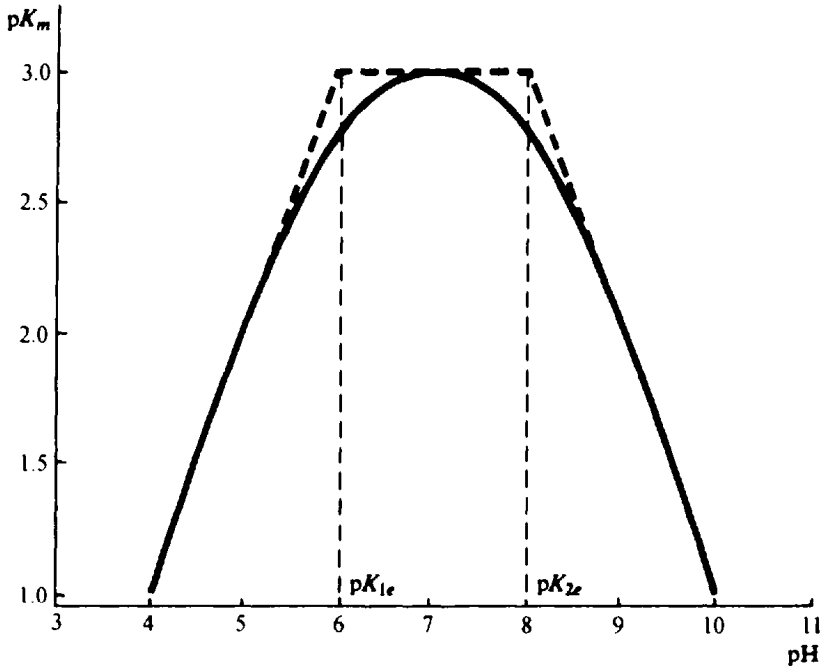


Fig. 9-3 A plot of $-\log$ (apparent $K_m \equiv pK_m$ versus pH for the enzyme scheme shown on page 261. The parameters used in Eq. (*9.35*) were $K_S = 10^{-3}$ mol L^{-1}, $K_{1e} = 10^{-6}$ mol L^{-1}, $K_{2e} = 10^{-8}$ mol L^{-1}.

pH has a slope of -1. In the intermediate pH range,when $K_{1e} \gg [H^+] \gg K_{2e}$, $K_m \simeq K_s$ and a plot of pK_m versus pH is a horizontal line. Hence the intersection of the horizontal line with the two straight-line segments yields estimates of pK_{1e} and pK_{2e} (Fig. 9-3).

9.10 MECHANISMS OF ENZYME INHIBITION

The equations given in Table 9.1 simply describe the inhibition behaviors of enzymes and thus can be called *phenomenological* expressions. However, it is important to describe basic mechanisms, in the way that Michaelis and Menten did for a single enzyme. The mechanisms account for the *form* of the inhibition equations.

Competitive Inhibition

The simplest scheme for competitive inhibition is

$$\begin{array}{ccccc} \mathrm{I} & & & & \\ + & & & & \\ \mathrm{E+S} & \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} & \mathrm{ES} & \xrightarrow{k_2} & \mathrm{E+P} \\ {\scriptstyle k_I}\Big\uparrow\Big\downarrow{\scriptstyle k_{-I}} & & & & \\ \mathrm{EI} & & & & \end{array}$$

where $K_I = k_{-I}/k_I$ and is the dissociation constant of the enzyme-inhibitor complex.

By using the steady-state analysis of Briggs and Haldane, it is possible to show that

$$v_0 = \frac{k_2[\mathrm{E}]_0[\mathrm{S}]_0}{\left(\dfrac{k_{-1}+k_2}{k_1}\right)\left(1+\dfrac{[\mathrm{I}]}{k_{-I}/k_I}\right)+[\mathrm{S}]_0} \tag{9.38}$$

$$= \frac{V_{\max}[\mathrm{S}]_0}{K_m\left(1+\dfrac{[\mathrm{I}]}{K_I}\right)+[\mathrm{S}]_0} \tag{9.39}$$

The double reciprocal form of Eq. (*9.39*) for competitive inhibition is:

$$\frac{1}{v_0} = \frac{1}{V_{\max}} + \frac{K_m}{V_{\max}}\left(1+\frac{[\mathrm{I}]}{K_I}\right)\frac{1}{[\mathrm{S}]_0} \tag{9.40}$$

This equation predicts that the slope of a Lineweaver-Burk plot will increase with increasing inhibitor concentration, but the intercept on the $1/v_0$ axis ($1/V_{\max}$) will not change. A series of plots for several experiments with different concentrations of inhibitor will all have the same $1/v_0$ intercept as shown in Fig. 9-4(*a*), indicating that competitive inhibition does not alter $V_{\max}$.

Noncompetitive Inhibition and Mixed Inhibition

The simplest scheme for noncompetitive and mixed inhibition is

$$\begin{array}{ccccc} \mathrm{I} & & \mathrm{I} & & \\ + & & + & & \\ \mathrm{E+S} & \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} & \mathrm{ES} & \xrightarrow{k_2} & \mathrm{E+P} \\ {\scriptstyle K_I}\Big\uparrow\Big\downarrow & & \Big\uparrow\Big\downarrow{\scriptstyle K_I'} & & \\ \mathrm{EI} & & \mathrm{ESI} & & \end{array}$$

The corresponding rate equation is

$$v_0 = \frac{V_{max}[S]_0}{K_m\left(1 + \frac{[I]}{K_I}\right) + \left(1 + \frac{[I]}{K_I'}\right)[S]_0} \tag{9.41}$$

This equation corresponds to that for mixed inhibition (Table 9.1) if $K_I \neq K_I'$. However, if $K_I = K_I'$, then pure noncompetitive inhibition is the result. Thus, it can be seen that, mechanistically speaking, pure noncompetitive inhibition is a special case of mixed inhibition.

The double reciprocal form of Eq. (9.41) for pure noncompetitive inhibition is:

$$\frac{1}{v_0} = \frac{(1 + [I]/K_I)}{V_{max}} + K_m \frac{(1 + [I]/K_I)}{V_{max}} \frac{1}{[S]_0} \tag{9.42}$$

This equation predicts that both the slope and the $1/v_0$ intercept of a Lineweaver-Burk plot will increase with increasing inhibitor concentration, but the intercept on the $1/[S]_0$ axis ($-1/K_m$) will not change. A series of plots for several experiments with different concentrations of inhibitor will all pass through the $1/[S]_0$ intercept as shown in Fig. 9-4(b), indicating that pure noncompetitive inhibition does not alter K_m.

Un- or Anticompetitive Inhibition

The simplest mechanism for anticompetitive inhibition is

$$\begin{array}{ccccc} & & & & I \\ & & & & + \\ E + S & \overset{k_1}{\underset{k_{-1}}{\rightleftharpoons}} & ES & \overset{k_2}{\longrightarrow} & E + P \\ & & \Big\| K_I & & \\ & & ESI & & \end{array}$$

and the corresponding rate equation is

$$v_0 = \frac{V_{max}[S_0]}{K_m + [S]_0(1 + [I]/K_I)} \tag{9.43}$$

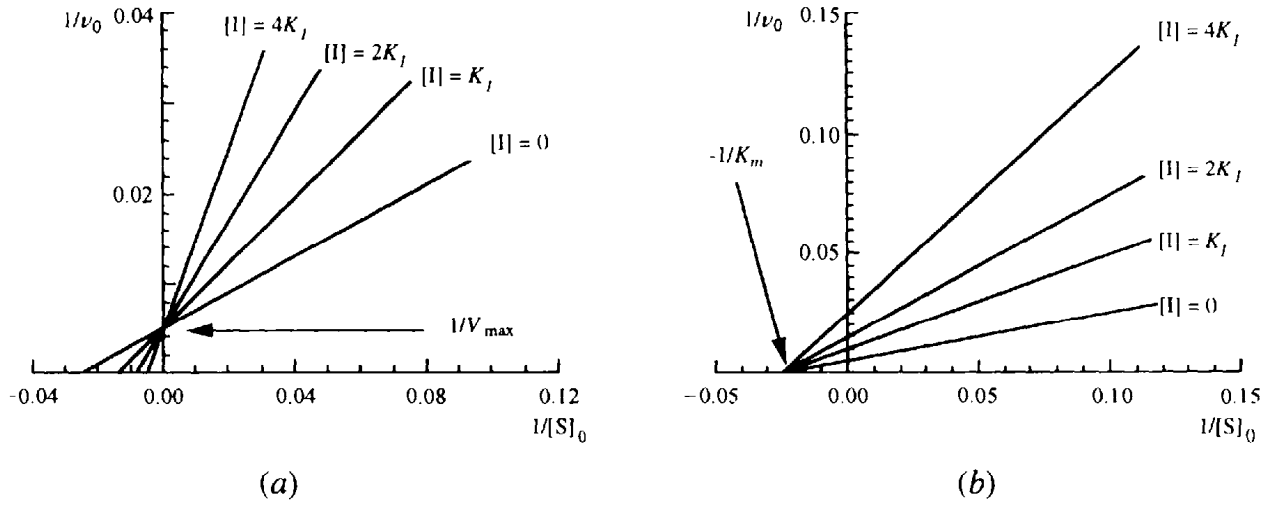


Fig. 9-4 Lineweaver-Burk plot: $1/v_0$ versus $1/[S]_0$ for (a) pure competitive inhibition and (b) pure noncompetitive inhibition.

In this case, Lineweaver-Burk plots of $1/v_0$ versus $1/[S]_0$ for several experiments with different concentrations of inhibitor yield parallel lines of slope K_m/V_{max}.

9.11 REGULATORY ENZYMES

General

The enzyme kinetics discussed thus far can be called *Michaelis-Menten* kinetics, since a plot of reaction rate versus substrate concentration is a pure rectangular hyperbola; alternatively, a plot of reciprocal initial velocity versus reciprocal substrate concentration is linear. Many reactions involving two or more substrates also give such linear plots with respect to one substrate, with the initial concentration of the other substrate(s) fixed. While many enzymes obey Michaelis-Menten kinetics, a significant number do not. These enzymes typically give velocity versus substrate curves that are *sigmoidal* rather than hyperbolic; they are called *control* or *regulatory* enzymes and are usually situated at the *beginning* or at *branch points* of a metabolic pathway.

EXAMPLE 9.7

The simplest form of regulation of a metabolic pathway is the *inhibition* of an enzyme by the product of the pathway. In Fig. 9-5, the E_i's denote enzymes, A and B are metabolites, and the circled minus sign indicates inhibition. If there were no inhibitor of the enzyme (E_1) acting on A, the concentration of B would depend entirely on its rate of synthesis or utilization. If the rate of utilization of B decreased or B was supplied from an outside source, its concentration would rise, perhaps even to toxic levels. However, if B is an *inhibitor* of the first enzyme, then as its concentration rises, the extent of inhibition will increase and its rate of synthesis will decrease. This effect is called *feedback inhibition* or negative feedback control; it is a concept also used in describing electronic circuits.

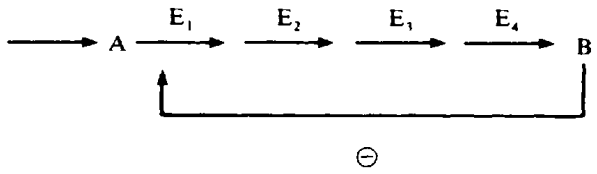


Fig. 9-5

EXAMPLE 9.8

Control in *branched* metabolic pathways is more complex. Consider the metabolic scheme in Fig. 9-6. Here, B reacts with C, and D is produced further along the pathway; for most effective control of D, B should inhibit the first enzyme (E_1) and C should activate it. In this case, if B is supplied from an external source so that $B \gg C$, then B would inhibit its own synthesis from A and the concentration of B and C would tend to become equal.

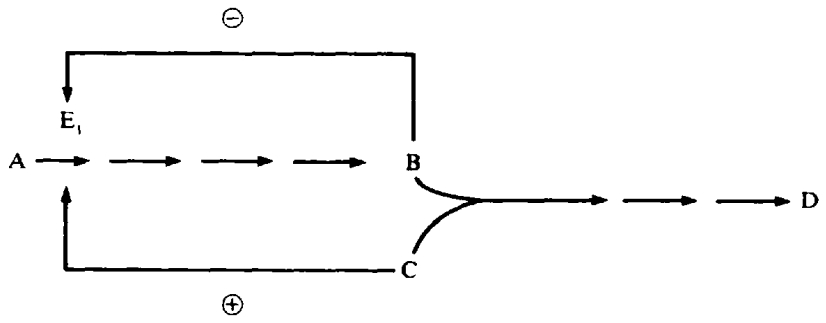
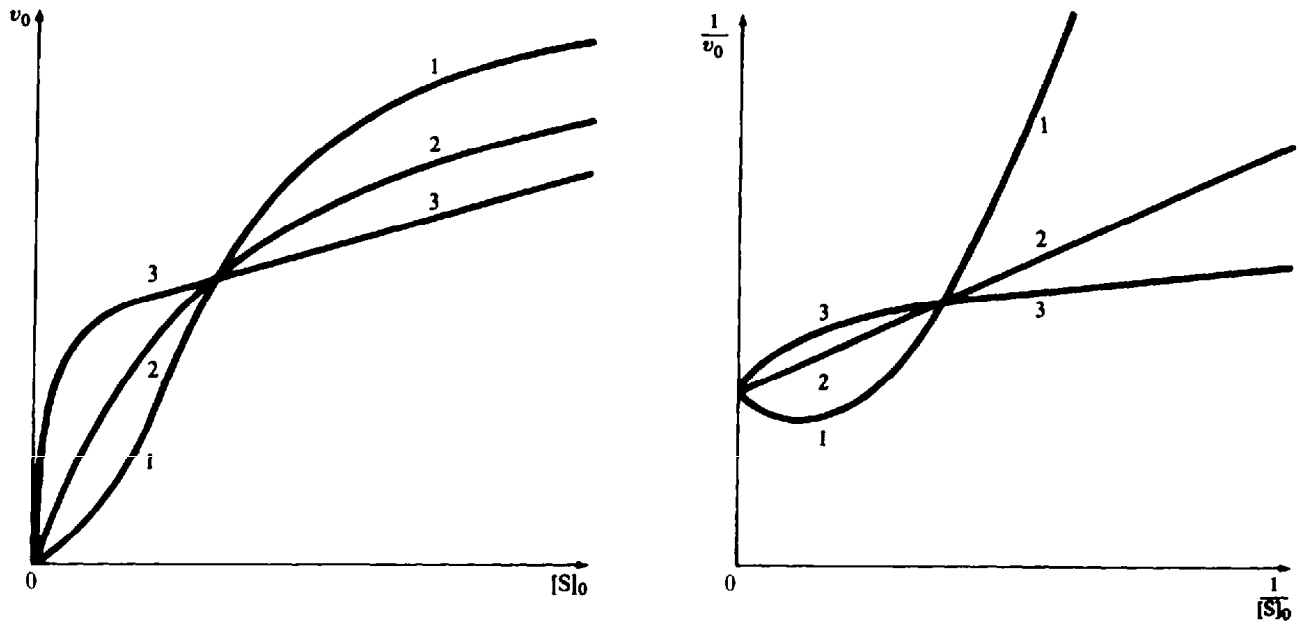


Fig. 9-6



(*a*) Plots of v_0 versus $[S]_0$: (1) sigmoidal; (2) hyperbolic; (3) apparent hyperbolic.

(*b*) The same data are plotted in double-reciprocal form. Note that only (2) is linear and therefore indicative of hyperbolic kinetics, while (1) is concave and (3) convex.

Fig. 9-7 Possible kinetic behavior of regulatory enzymes.

Alternatively, if $C \gg B$, then C would activate the production of B and this again would tend to equalize the concentrations of B and C. This activation by C is usually the result of C competing for the same binding site as B on E_1, and thus reducing the inhibition by B. The first enzyme of pyrimidine synthesis, aspartate carbamoyltransferase, in *E. coli*, is subject to this type of control (Example 8.5; Chap. 15); in this case B is CTP, C is ATP, and D is the nucleic acids.

Kinetic Behavior of Regulatory Enzymes

Usually effector molecules bear little structural resemblance to the substrates of the enzymes they control. The control is therefore not likely to be due to binding at the active site but at an alternative site, the *allosteric site*. The effect on the reaction at the active site is mediated by conformational changes in the protein.

If an effector of an enzyme is also the substrate, it is called a *homotropic* effector; if it is a nonsubstrate, it is called *heterotropic*.

Regulatory enzymes are usually identified by the deviation of their kinetics from Michaelis-Menten kinetics; plots of velocity versus substrate concentration can be a sigmoidal curve or a modified hyperbola [Fig. 9-7(*a*)]. If these curves are plotted in the double-reciprocal (Lineweaver-Burk) form, nonlinear graphs are obtained [Fig. 9-7(*b*)].

A useful parameter for comparing regulatory enzymes is the ratio R_s:

$$R_s = \frac{\text{substrate concentration at 0.9 maximal velocity}}{\text{substrate concentration at 0.1 maximal velocity}} \tag{9.44}$$

For a Michaelis-Menten enzyme $R_s = 81$. For a sigmoidal curve, $R_s < 81$ and the enzyme is said to exhibit *positive cooperativity* with respect to the substrate. Positive cooperativity implies that the substrate binding or catalytic rate, or both, increases with increasing substrate concentration more than would be expected for a simple Michaelis-Menten enzyme. If $R_s > 81$, the enzyme is said to display *negative cooperativity* with respect to the substrate; substrate binding or catalysis, while increasing, becomes progressively less than would be found with a simple Michaelis-Menten enzyme as substrate concentration is increased.

EXAMPLE 9.9

Hetero- and homotropic effects can also operate with binding proteins. The best example is the binding of O_2 to hemoglobin; the binding of one molecule of O_2 to hemoglobin *cooperates* in a *positive* way with the binding of the next molecule, so that the apparent affinity of hemoglobin for O_2 increases as the *degree of saturation* by O_2 increases. When the fractional saturation is plotted against the partial pressure of O_2 (equivalent to O_2 concentration), the curve is not hyperbolic but *sigmoidal* (Fig. 9-8).

In human red blood cells, the *heterotropic effector* 2,3-bisphosphoglycerate (BPG) reduces the affinity of all four binding sites of the hemoglobin tetramer for O_2 (Chap. 5). On the other hand, myoglobin, the O_2-binding protein of muscle, is a single subunit and does *not* display sigmoidal O_2-binding behavior, nor is it affected by BPG.

Mathematical Models of Cooperativity

Although hemoglobin is not an enzyme, but an O_2-binding protein, the study of it has contributed a great deal to our understanding of molecular cooperativity. The cooperative O_2-binding behavior of hemoglobin was first recognized by Bohr (1903), long before the effect was seen with enzymes; much effort was expended in developing theories to explain the effect. Therefore, it is worth considering some of the earlier theories, which still have considerable relevance today.

The Hill Equation

In 1909 A. V. Hill proposed that the binding reaction between hemoglobin (Hb) and oxygen could be described by a reaction of molecularity $n + 1$:

$$\mathrm{Hb} + n\mathrm{O_2} \rightleftharpoons \mathrm{Hb(O_2)}_n$$

where n is the number of O_2-binding sites on a hemoglobin molecule. The equation for the fractional saturation Y (the fraction of binding sites occupied by O_2 at any instant) is (Problem 5.21),

$$Y = \frac{K_b(pO_2)^n}{1 + K_b(pO_2)^n} \tag{9.45}$$

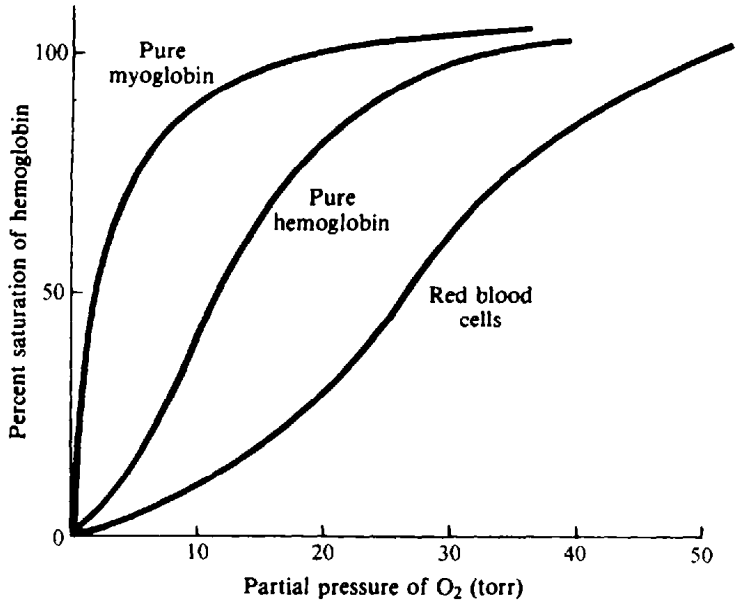


Fig. 9-8 Oxygen binding curves for hemoglobin and myoglobin.

where pO_2 is the partial pressure of O_2 (torr) and K_b is the *association binding constant*, defined as

$$K_b = \frac{[Hb(O_2)_n]}{[Hb](pO_2)^n} \tag{9.46}$$

Note that Y is a dimensionless quantity that has the range 0–1; n is called the *Hill coefficient*. For values of n around 2.5, a sigmoidal curve akin to those in Fig. 9-8 was obtained by Hill. The fact that n is not an integer raised some early problems in the mechanistic interpretations of the data analyzed with the Hill equation. This enigma led to the development of other mechanistic models.

EXAMPLE 9.10

If $n = 1$ in the Hill equation, what form does the equation take?

When $n = 1$, Eq. (*9.45*) yields a rectangular hyperbola. Recall that the Michaelis-Menten expression is written with K_m being equivalent to (and having the units of) a *dissociation* constant, whereas binding equations, such as the Hill equation, are usually written with *association* constants (their numerical values being the reciprocal of the corresponding dissociation constants).

Question: How is it possible to determine the value of K_b and n in the Hill equation from experimental data?

By rearranging Eq. (*9.46*) and taking the logarithm of both sides, we obtain

$$\log\left(\frac{[Hb(O_2)_n]}{[Hb]}\right) = \log K_b + n\log(pO_2) \tag{9.47}$$

Or, a more general form of the expression using the fractional saturation Y of a binding protein with substrate S is

$$\log\left(\frac{Y}{1-Y}\right) = \log K_b + n\log[S]_0 \tag{9.48}$$

A plot of the left-hand side of Eq. (*9.48*) versus $\log[S]_0$ is called a *Hill plot*; it yields an estimate of n from the slope, and K_b from the ordinate intercept.

EXAMPLE 9.11

For an *enzymatic* reaction (in contrast to a *binding* reaction), the initial velocity v_0 is determined by the concentration of the enzyme substrate complex; therefore, what form does equation (*9.48*) take?

A fractional saturation of unity corresponds to V_{max}, and thus the Hill plot consists of $\log(v_0/V_{max} - v_0)$ versus $\log[S]_0$. Clearly, an initial estimate of V_{max} must be made before analyzing the data this way.

The Adair Equation

G. S. Adair in 1925 deduced that the molecular weight of hemoglobin was about four times as great as previously thought. He postulated that hemoglobin had four O_2-binding sites that were filled in a four-step process as follows:

$$\begin{aligned} Hb + O_2 &\rightleftharpoons HbO_2 \\ HbO_2 + O_2 &\rightleftharpoons Hb(O_2)_2 \\ Hb(O_2)_2 + O_2 &\rightleftharpoons Hb(O_2)_3 \\ Hb(O_2)_3 + O_2 &\rightleftharpoons Hb(O_2)_4 \end{aligned}$$

Consider *one* binding site on a totally empty (no sites filled) hemoglobin molecule; the binding

reaction for this site can be characterized by the *association* equilibrium constant $K_1 = k_1/k_{-1}$, where k_1 and k_{-1} are the unitary rate constants for the forward and reverse reactions, respectively. This equilibrium constant is called an *intrinsic* constant since it refers to one site only. The *overall*, or *extrinsic*, binding constant relates to all four of the binding sites and thus has the value $4K_1$. The extrinsic binding constants for the other three sites are also expressed in terms of *intrinsic* binding constants and their so-called *statistical factors*: $\frac{3}{2}K_2$, $\frac{2}{3}K_3$, and $\frac{1}{4}K_4$. Thus, the Adair binding function Y for four sites is expressed as follows:

$$Y = \frac{\text{number (or moles) of binding sites occupied}}{\text{total number (or moles) of binding sites}} \tag{9.49}$$

$$= \frac{K_1[\mathrm{X}] + 3K_1K_2[\mathrm{X}]^2 + 3K_1K_2K_3[\mathrm{X}]^3 + K_1K_2K_3K_4[\mathrm{X}]^4}{1 + 4K_1[\mathrm{X}] + 6K_1K_2[\mathrm{X}]^2 + 4K_1K_2K_3[\mathrm{X}]^3 + K_1K_2K_3K_4[\mathrm{X}]^4} \tag{9.50}$$

where $[\mathrm{X}] = pO_2$ in the particular case of hemoglobin binding.

EXAMPLE 9.12

An alternative way of viewing the concept of extrinsic binding constants is as follows. The experimentally *measured* binding constant for each site of a polymeric protein, or enzyme, will depend on the number of available sites on each molecule; e.g., for the hemoglobin tetramer the first binding reaction is

$$\mathrm{Hb} + \mathrm{O_2} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{HbO_2}$$

There are four sites available for binding O_2, but only one from which bound O_2 can dissociate. Thus, from the law of mass action, the (overall) extrinsic equilibrium constant is equal to $4k_1/k_{-1} = 4K_1$. Similarly, the *extrinsic* constant denoted by K_2^e is given by $\frac{3}{2}K_2$, where K_2 is the intrinsic constant, and by the same reasoning $K_3^e = \frac{2}{3}K_3$ and $K_4^e = \frac{1}{4}K_4$.

Question: If the hemoglobin tetramer had four *identical noninteracting* binding sites, what relative values would the intrinsic binding constants have?

They would all be equal, i.e., $K_1 = K_2 = K_3 = K_4$. There would thus be no cooperativity in this case (see Example 9.13). If each binding step facilitated the next, i.e., $K_1 < K_2 < K_3 < K_4$, then positive cooperativity would exist. Negative cooperativity at each stage in the binding requires that $K_1 > K_2 > K_3 > K_4$. Clearly, more complex relationships between the values of the K's *could* exist; e.g., $K_1 > K_2 < K_3 < K_4$.

EXAMPLE 9.13

Show that if $K_1 = K_2 = K_3 = K_4$, the four-site Adair equation becomes that of a rectangular hyperbola.

Note that the statistical factors in the numerator and denominator of the four-site Adair equation, and in fact in the general n-site equation, are the coefficients of the *binomial expansion*, i.e., the coefficients of the expansion of $(1 - x)^n$. In the present case with $n = 4$, Eq. (*9.50*) applies and

$$Y = \frac{K_1[\mathrm{X}] + 3K_1^2[\mathrm{X}]^2 + 3K_1^3[\mathrm{X}]^3 + K_1^4[\mathrm{X}]^4}{1 + 4K_1[\mathrm{X}] + 6K_1^2[\mathrm{X}]^2 + 4K_1^3[\mathrm{X}]^3 + K_1^4[\mathrm{X}]^4} \tag{9.51}$$

$$= \frac{K_1[\mathrm{X}](1 + K_1[\mathrm{X}])^3}{(1 + K_1[\mathrm{X}])^4} = \frac{K_1[\mathrm{X}]}{1 + K_1[\mathrm{X}]} \tag{9.52}$$

The final expression is that of a hyperbola; it is evident that it is equivalent to the Michaelis-Menten expression when it is recalled that K_1 is an *association* constant.

The Model of Monod, Wyman, and Changeux (MWC Model) for Homotropic Effects

The cooperative binding of O_2 by hemoglobin and the allosteric effects in many enzymes require interaction between sites that are widely separated in space. The MWC model was proposed in 1965 to incorporate allosteric and conformational effects in an explanation of enzyme cooperativity. The seminal observation was that most cooperative proteins have several identical subunits (*protomers*) in each molecule (*oligomer*); this situation is imperative for binding cooperativity. The MWC model is defined as follows:

1. Each protomer can exist in either of two conformations, designated R and T; these originally referred to *relaxed* and *tense*, with the latter not being catalytically active.
2. All subunits of one oligomer *must* occupy the *same* conformation; hence, for a tetrameric protein, the conformational states R_4 and T_4 are the only ones permitted. Mixed states such as R_3T are forbidden. In other words, a *concerted* (concert: "all together") transition takes place in the conversion of R_4 to T_4 and vice versa.
3. The two states of the protein are in equilibrium *independent* of whether any ligand (substrate) is bound; for a tetramer the equilibrium constant is

$$L = [T_4]/[R_4] \tag{9.53}$$

4. A ligand molecule can bind to a subunit in either conformation, but the association constants are different. For each R subunit,

$$K_R = [RX]/[R][X] \tag{9.54}$$

and for each T subunit,

$$K_T = [TX]/[T][X] \tag{9.55}$$

These postulates imply the following multiple equilibrium scheme for the protein and ligand (X):

$$\begin{array}{ccc}
 & \overset{L}{\rightleftharpoons} & \\
R_4 & & T_4 \\
{\scriptstyle 4K_R[X]}\uparrow\downarrow & & \uparrow\downarrow{\scriptstyle 4K_T[X]} \\
R_4X & \rightleftharpoons & T_4X \\
{\scriptstyle \frac{3}{2}K_R[X]}\uparrow\downarrow & & \uparrow\downarrow{\scriptstyle \frac{3}{2}K_T[X]} \\
R_4X_2 & \rightleftharpoons & T_4X_2 \\
{\scriptstyle \frac{2}{3}K_R[X]}\uparrow\downarrow & & \uparrow\downarrow{\scriptstyle \frac{2}{3}K_T[X]} \\
R_4X_3 & \rightleftharpoons & T_4X_3 \\
{\scriptstyle \frac{1}{4}K_R[X]}\uparrow\downarrow & & \uparrow\downarrow{\scriptstyle \frac{1}{4}K_T[X]} \\
R_4X_4 & \rightleftharpoons & T_4X_4
\end{array}$$

The concentrations of the 10 forms of the protein are related by the following equilibrium expressions, in which c is the ratio of K_T/K_R.

$$\begin{aligned}
[R_4X] &= 4K_R[R_4][X] \\
[R_4X_2] &= \tfrac{3}{2}K_R[R_4X][X] = 6K_R^2[R_4][X]^2 \\
[R_4X_3] &= \tfrac{2}{3}K_R[R_4X_2][X] = 4K_R^3[R_4][X]^3 \\
[R_4X_4] &= \tfrac{1}{4}K_R[R_4X_3][X] = K_R^4[R_4[X]^4 \\
[T_4] &= L[R_4] \\
[T_4X] &= 4K_T[T_4][X] = 4LcK_R[R_4][X] \\
[T_4X_2] &= 6K_T^2[T_4][X]^2 = 6Lc^2K_R^2[R_4][X]^2
\end{aligned} \tag{9.56}$$

$$[T_4X_3] = 4K_T^3[T_4][X]^3 \quad = 4Lc^3K_R^3[R_4][X]^3$$
$$[T_4X_4] = K_T^4[T_4][X]^4 \quad = Lc^4K_R^4[R_4][X]^4$$

Again, the statistical factors 4, $\frac{3}{2}$, etc., arise because K_R and K_T are *intrinsic* association binding constants, yet the overall expressions for the complexes requires *extrinsic* parameters; e.g., $K_R = [RX]/[R][X] = \frac{3}{2}[R_4X_2]/[R_4X][X]$ because there are three unfilled sites in the R_4X molecule and two in each R_4X_2 molecule (see also Example 9.12). The fractional saturation Y is again defined by Eq. (*9.49*):

$$Y = \frac{[R_4X] + 2[R_4X_2] + 3[R_4X_3] + 4[R_4X_4] + [T_4X] + 2[T_4X_2] + 3[T_4X_3] + 4[T_4X_4]}{4([R_4] + [R_4X] + [R_4X_2] + [R_4X_3] + [R_4X_4] + [T_4] + [T_4X] + [T_4X_2] + [T_4X_3] + [T_4X_4])} \tag{9.57}$$

$$= \frac{(1 + K_R[X])^3K_R[X] + Lc(1 + cK_R[X])^3K_R[X]}{(1 + K_R[X])^4 + L(1 + cK_R[X])^4} \tag{9.58}$$

The general n-site MWC model with $K_R[X] = \alpha$ is

$$Y = \frac{\alpha(1 + \alpha)^{n-1} + Lc\alpha(1 + c\alpha)^{n-1}}{(1 + \alpha)^n + L(1 + c\alpha)^n} \tag{9.59}$$

The behavior of this expression for various ranges of parameter values is shown in Fig. 9-9.

EXAMPLE 9.14

The shape of the saturation curve defined by Eq. (*9.59*) depends on the values of L and c. If $L = 0$, then the T form of the protein does not exist and $Y = K_R[X]/(1 + K_R[X])$. This defines a hyperbolic binding function. Similarly if $L = \infty$, $Y = K_T[X]/(1 + K_T[X])$. Thus, deviations from hyperbolic binding occur only if both R and T forms exist; otherwise the situation described for the Adair equation in Example 9.13 applies since binding is *independent* and *identical* at each site.

EXAMPLE 9.15

The fact that Eq. (*9.59*), or (*9.58*), defines sigmoidal curves is not obvious unless $c = 0$. In this case, for $n = 4$,

$$Y = \frac{(1 + K_R[X])^3K_R[X]}{L + (1 + K_R[X])^4} \tag{9.60}$$

Hence, when [X] is large and $K_R[X] \gg L$, the term in L is negligible and $Y \approx \dfrac{K_R[X]}{1 + K_R[X]}$; i.e., the hyperbolic binding equation is the result. At *low* values of [X], the term in L dominates the denominator and the slope of the binding curve as $[X] \to 0$ approaches $\dfrac{K_R}{(L + 1)}$, in contrast to the slope of the hyperbola, which approaches K_R. In symbols,

$$\lim_{[X]\to 0} \frac{dY}{d[X]} = \frac{K_R}{(L + 1)} \tag{9.61}$$

and the curve *must* be sigmoidal if L is large compared with 1; in the limit of infinite L, the initial slope is zero.

The MWC Model for Heterotropic Effects

The MWC model of the previous section accounts only for *homotropic* effects, as occur, for example, with O_2 binding to hemoglobin, but *heterotropic* effects are frequently observed with regulatory enzymes. To incorporate these effects into the MWC model, we must introduce the binding

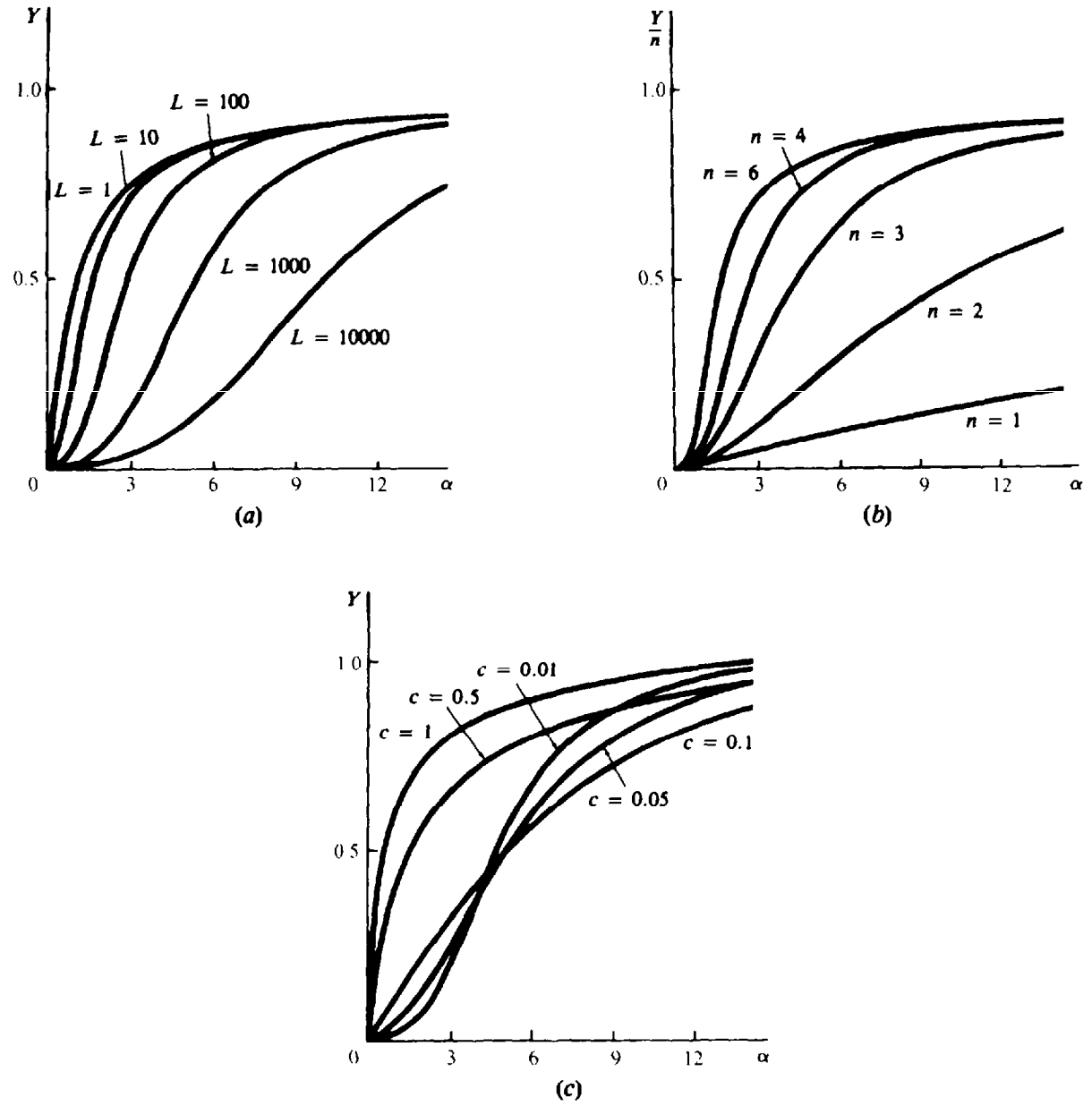


Fig. 9-9 Binding curves described by the MWC equation Eq. (*9.59*). (*a*) The effect of varying the value of the isomerization constant (L) with $n = 4$ and $c = 0.01$. (*b*) The effect of varying the number of binding sites (n) with $L = 100$ and $c = 0.01$. (*c*) The effect of varying c, the ratio of K_T/K_R, with $L = 1{,}000$ and $n = 4$.

of a second ligand to either the R or the T state, or to both. Let us consider a simplified model and thus assume that the substrate X (ligand) binds only to the R state. For the binding of another ligand, A, to bring about activation of the enzyme, it must promote a shift of the equilibrium between the R and T states in favor of the R state. An activator must therefore bind to the R state. By the same reasoning an inhibitor, I, must bind to the T state and thus cause a shift of the enzyme into this inactive state. By a multiple equilibrium analysis similar to that used in the previous subsection the expression for Y is obtained:

$$Y = \frac{\alpha(1+\alpha)^{n-1}}{(1+\alpha)^n + [L(1+\beta)^n/(1+\gamma)^n]} \tag{9.62}$$

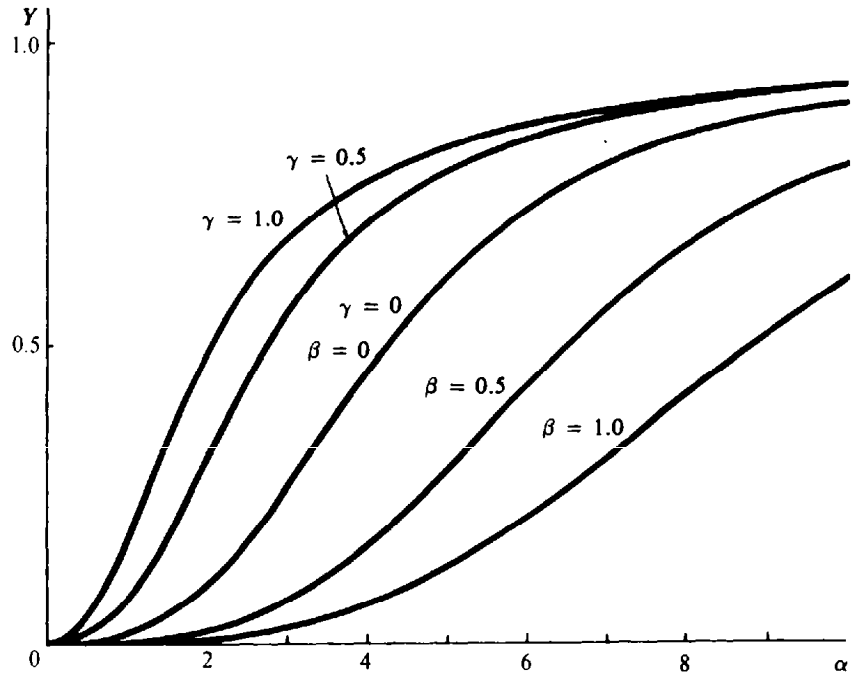


Fig. 9-10 Behavior of an MWC allosteric enzyme in the presence of positive and negative heterotropic effectors. The activator term, γ, in Eq. (*9.62*) causes the curve to become more hyperbolic, whereas the inhibitor term (β) renders it more sigmoidal. The curves were constructed using Eq. (*9.62*) with $L = 1{,}000$ and $n = 4$.

where $\alpha = K_R[X]$, $\beta = K_I[I]$, $\gamma = K_A[A]$, and $L = [T_n]/[R_n]$. The term $L(1+\beta)^n/(1+\gamma)^n$ is regarded as an *allosteric coefficient*; if this term is zero, then Eq. (*9.62*) reverts to a hyperbolic binding function. Clearly, an increase in the activator term, γ, *decreases* the value of the allosteric coefficient, thus making the function more hyperbolic (see Example 9.15). An increase of inhibitor concentration increases the inhibition term (β), thus increasing the allosteric coefficient and making the curve more sigmoidal. Both these effects are seen in Fig. 9-10.

The Model of Koshland, Nemethy, and Filmer (KNF Model)

The failure of the MWC model to describe negatively cooperative hetero- and homotropic effects in the binding function of an oligomeric protein led Koshland et al. (1966) to develop a more general model of cooperativity. Koshland's 1958 *induced-fit hypothesis* for enzyme specificity extended Fischer's *lock-and-key* concept. Koshland claimed that the binding of substrate to an enzyme *creates* the correct three-dimensional arrangement of reactive groups for catalysis to occur. This abstract notion was extended in the KNF model of oligomeric cooperative enzymes; in this case, the change in conformation is *induced* (in contrast to the MWC model, where equilibrium between states of the subunits exists whether substrate is bound or not) by the binding of the substrate to a protomer (subunit). A change in the conformation in the protomer is transmitted to neighboring protomers to affect their binding and catalytic properties. This basic idea can explain both negative *and* positive cooperativity. However, the details of the model will not be given here.

Solved Problems

INTRODUCTION AND DEFINITIONS

9.1. Use the principle of mass action to write expressions for the rate equations for the following reactions:

(*a*) $A \xrightarrow{k_1} P$

(*b*) $A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P$

(*c*) $A + B \xrightarrow{k_1} P$

(*d*) $A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P$

(*e*) $A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} P + Q + R$

SOLUTION

(*a*) $\dfrac{d[A]}{dt} = -k_1[A] \qquad \dfrac{d[P]}{dt} = -\dfrac{d[A]}{dt}$

(*b*) $\dfrac{d[A]}{dt} = -k_1[A] + k_{-1}[P] \qquad \dfrac{d[P]}{dt} = -\dfrac{d[A]}{dt}$

(*c*) $\dfrac{d[A]}{dt} = \dfrac{d[B]}{dt} = -k_1[A][B] \qquad \dfrac{d[P]}{dt} = -\dfrac{d[A]}{dt}$

(*d*) $\dfrac{d[A]}{dt} = \dfrac{d[B]}{dt} = -k_1[A][B] + k_{-1}[P] \qquad \dfrac{d[P]}{dt} = -\dfrac{d[A]}{dt}$

(*e*) $\dfrac{d[A]}{dt} = \dfrac{d[B]}{dt} = -k_1[A][B] + k_{-1}[P][Q][R] \qquad \dfrac{d[P]}{dt} = \dfrac{d[Q]}{dt} = \dfrac{d[R]}{dt} = -\dfrac{d[A]}{dt}$

9.2. What is the kinetic order of each of the reactions in Prob. 9.1?

SOLUTION

Reactions (*a*) and (*b*) are first-order, in each direction for (*b*); (*c*) and (*d*) are second-order in the *forward* direction (left to right), and first-order in the reverse direction for (*d*); (*e*) is second-order in the forward direction and third-order in the reverse direction.

Note: For the uncatalyzed reactions, the *molecularity* and the *order* have been given as being the same. However, a trimolecular reaction, as depicted in (*e*), is an unlikely occurrence; what is more likely is the interaction between two of the reactants, *followed* by reaction with the third. This amounts to two second-order reactions in sequence.

9.3. What are the *units* of the *rate constants* in the reaction schemes of Prob. 9.1?

SOLUTION

(*a*) k_1: s^{-1}; (*b*) k_1 and k_{-1}: s^{-1}; (*c*) k_1: $mol^{-1}\,L\,s^{-1}$; (*d*) k_1: $mol^{-1}\,L\,s^{-1}$, and k_{-1}: s^{-1}; (*e*) k_1: $mol^{-1}\,L\,s^{-1}$, and k_{-1}: $mol^{-2}\,L^2\,s^{-1}$.

9.4. The data in Table 9.2 were obtained for the *rate* of a reaction that has the stoichiometry $aA + bB \rightarrow pP$; the rate was determined with various concentrations of A and B. Determine the order of the reaction; i.e., estimate the values of a and b.

SOLUTION

The order of a reaction is an *experimentally* determined parameter, and the simplest way of obtaining it is to measure the initial reaction rate for a range of concentrations of reactants. Then a plot of log(rate)

Table 9.2. Reaction Order Data: v_0 (μmol $L^{-1} s^{-1}$)

		$[A]_0$ (mmol L^{-1})			
		10	20	50	100
$[B]_0$ (mmol L^{-1})	10	1.2	2.0	2.8	3.9
	20	2.6	4.0	5.9	7.9
	50	6.5	8.8	14.5	19.8
	100	12.5	17.9	29.0	40.5

versus log(concentration) of one of the reactants, while holding the rest constant, gives a line with a slope equal to the order. If the concentrations of all reactants are varied in a constant ratio, then the slope of the graph is the *overall* reaction order; if just one is varied, the order refers only to *that* reactant. Hence, from the data in Table 9.2, a plot of log(diagonal elements) versus log(concentration) yields Fig. 9-11(*a*).

The slope in Fig. 9-11(*a*) is $1\frac{1}{2}$, which is therefore the overall order of the reaction. A family of four order plots is obtained for variable $[A]_0$ and constant $[B]_0$ *and* vice versa [Fig. 9-11(*b*) and (*c*)]. From the slopes of these plots, we see that a, the order of A, is $\frac{1}{2}$ and b, the order of B, is 1.

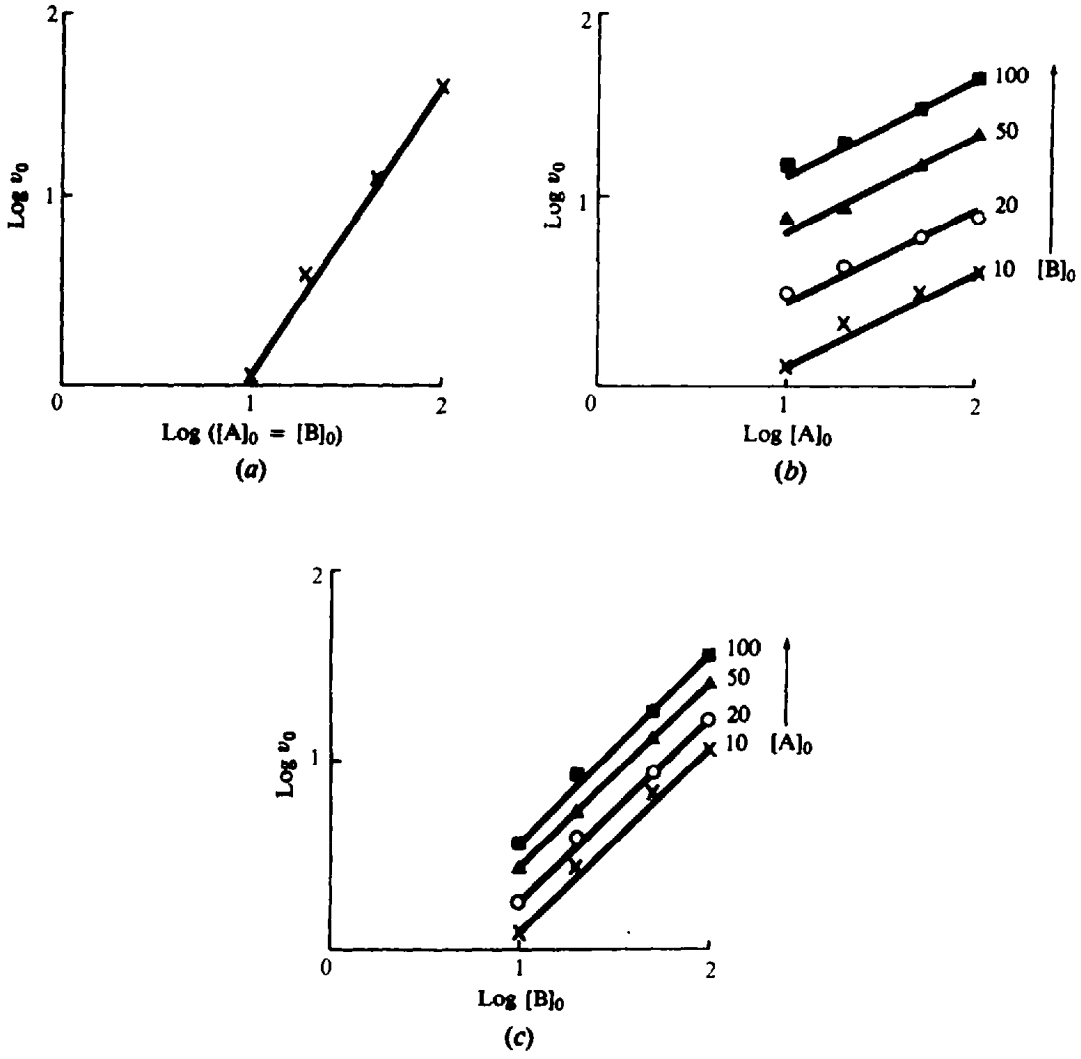


Fig. 9-11 Reaction order plots.

DEPENDENCE OF ENZYME REACTION RATE ON SUBSTRATE CONCENTRATION

9.5. Use the Michaelis-Menten equation to complete the enzyme kinetic data set; the K_m is known to be 1 mmol L^{-1}.

$[S]_0$ (mmol L^{-1})	v_0 (μmol L^{-1} min^{-1})
0.5	50
1.0	—
2.0	—
3.0	—
10.0	—

SOLUTION

Using Eq. (9.5) and the first entry in the table gives $V_{max} = 150\ \mu$mol L^{-1} min^{-1}. The other entries simply follow by substituting the $[S]_0$ values into Eq. (9.5); the results are shown below.

$[S]_0$ (mmol L^{-1})	v_0 (μmol L^{-1} min^{-1})
0.5	50
1.0	75
2.0	100
3.0	112.5
10.0	136.4

9.6. Hexokinase catalyzes the phosphorylation of glucose and fructose by ATP. However the K_m for glucose is 0.13 mmol L^{-1}, whereas that for fructose is 1.3 mmol L^{-1}. Assume V_{max} is the same for both glucose and fructose and the enzyme displays *hyperbolic kinetics* [Eq. (9.5)]. (*a*) Calculate the normalized initial velocity of the reaction (i.e., v_0/V_{max}) for each substrate when $[S]_0 = 0.13$, 1.3, and 13.0 mmol L^{-1}. (*b*) For which substrate does hexokinase have the greater affinity?

SOLUTION

(*a*) For glucose the values of v_0/V_{max} are 0.5, 0.91, and 0.99; for fructose the values are 0.091, 0.5, and 0.91.

(*b*) Glucose; at lower concentrations, the reaction rate is a greater fraction of V_{max} than it is with fructose.

GRAPHICAL EVALUATION OF K_m AND V_{max}

9.7. A constant amount of enzyme solution was added to a series of reaction mixtures containing different substrate concentrations. The initial reaction rates were obtained by measuring the initial slopes of the progress curves of product formation. The data in Table 9.3 were obtained:

(*a*) What is the V_{max} for this enzyme-reaction mixture?

(*b*) What is the K_m of the enzyme for the substrate?

Table 9.3. Steady-State Enzyme Kinetic Data

$[S]_0$ (μmol L^{-1})	v_0 (μmol L^{-1} min^{-1})
0.1	0.27
2.0	5.0
10.0	20
20.0	40
40.0	64
60.0	80
100.0	100
200.0	120
1,000.0	150
2,000.0	155

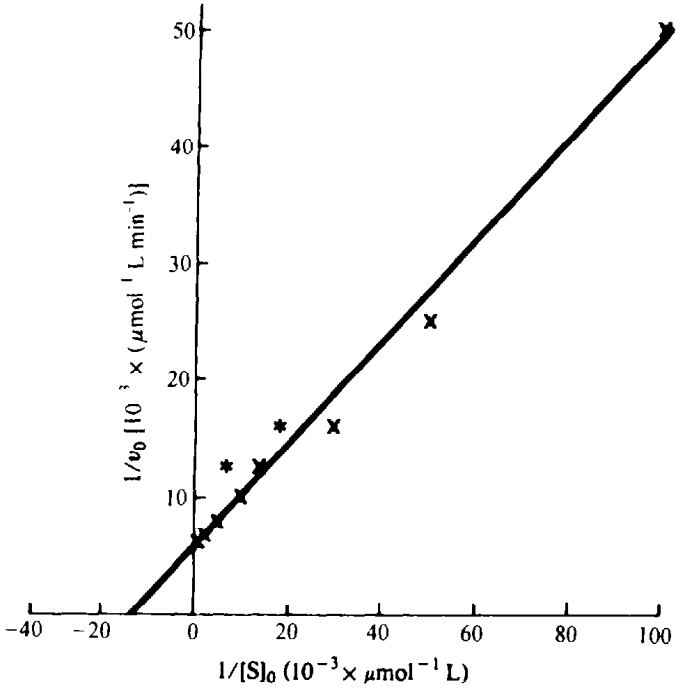


Fig. 9-12 Lineweaver-Burk plot of all but the first two data pairs in Table 9.3. (The asterisks refer to Prob. 9.9.)

SOLUTION

We can use the Lineweaver-Burk equation; for this, the reciprocals of the variables in Table 9.3 must be calculated and then plotted as shown in Fig. 9-12. (Ignore the asterisks in Fig. 9-12; they refer to Prob. 9.9.) (*a*) From the reciprocal of the ordinate intercept, $V_{max} = 160\,\mu$mol L^{-1} min^{-1}, and (*b*) from the reciprocal of the abscissal intercept, $K_m = 60\,\mu$mol L^{-1}.

9.8. One of the critical factors in experimental design for enzyme kinetics is the correct choice of substrate concentrations. On the basis of the data in Fig. 9-12, what is the possible optimal substrate concentration range?

SOLUTION

The most useful concentrations are those around K_m. In fact, a useful rule of thumb is 1/5 to 5 K_m; this way, v_0 will range from 0.17 to 0.83 times V_{max}. In other words, v_0 will be ~5 times greater at the highest concentration than at the lowest. If initial concentrations are too low, then velocity estimates become less accurate because of the nonlinearity of the reaction-progress curves near time $t = 0$.

9.9. For the enzyme kinetic experiment described in Prob. 9.7, an additional two data pairs were obtained: ($[S]_0 = 50\ \mu mol\ L^{-1}$, $v_0 = 60\ \mu mol\ L^{-1}\ min^{-1}$) and ($[S]_0 = 150\ \mu mol\ L^{-1}$, $v_0 = 80\ \mu mol\ L^{-1}\ min^{-1}$). Construct a Lineweaver-Burk plot from the whole data set. What are the estimates of K_m and V_{max} now, as distinct from the results obtained in Prob. 9.7?

SOLUTION

The points marked by an asterisk in Fig. 9-12 are the new ones. They lead to a much lower estimated value of V_{max} and a lower value of K_m, approximately $25\ \mu mol\ L^{-1}\ min^{-1}$ and $10\ \mu mol\ L^{-1}$, respectively.

ENZYME INHIBITION

9.10. The effect of an inhibitor I on the rate of a single-substrate enzyme-catalyzed reaction was investigated and gave the following results:

Substrate Concentration $[S]_0$ ($mmol\ L^{-1}$)	Inhibitor Concentration ($mmol\ L^{-1}$)		
	0	0.5	1.0
	Rate of Reaction v_0 ($\mu mol\ L^{-1}\ min^{-1}$)		
0.05	0.33	0.20	0.14
0.10	0.50	0.33	0.25
0.20	0.67	0.50	0.40
0.40	0.80	0.67	0.57
0.50	0.83	0.71	0.63

(*a*) What is the mode of action of the inhibitor? (*b*) Estimate values for V_{max}, K_m, and the inhibition constant(s) for the reaction.

SOLUTION

(*a*) The simplest way to deduce the mode of action of the inhibitor is to plot $1/v_0$ as a function of $1/[S]_0$ for each inhibitor concentration (Fig. 9-13).

Equation (*9.40*) for pure competitive inhibition predicts that the slope of a Lineweaver-Burk plot will change with changes in inhibitor concentration, but the intercept on the $1/v_0$ axis ($1/V_{max}$) will not. Thus, the above data are consistent with the inhibitor (I) acting as a *pure competitive* inhibitor.

(*b*) From the intercept on the $1/v_0$ axis, V_{max} may be estimated. We note from the equation above that the slope (Sl) of the plot is:

$$Sl = \frac{K_m}{V_{max}}\left(1 + \frac{[I]}{K_I}\right)$$

Thus, when the slope is plotted as a function of [I], we get the results shown in Fig. 9-14. The

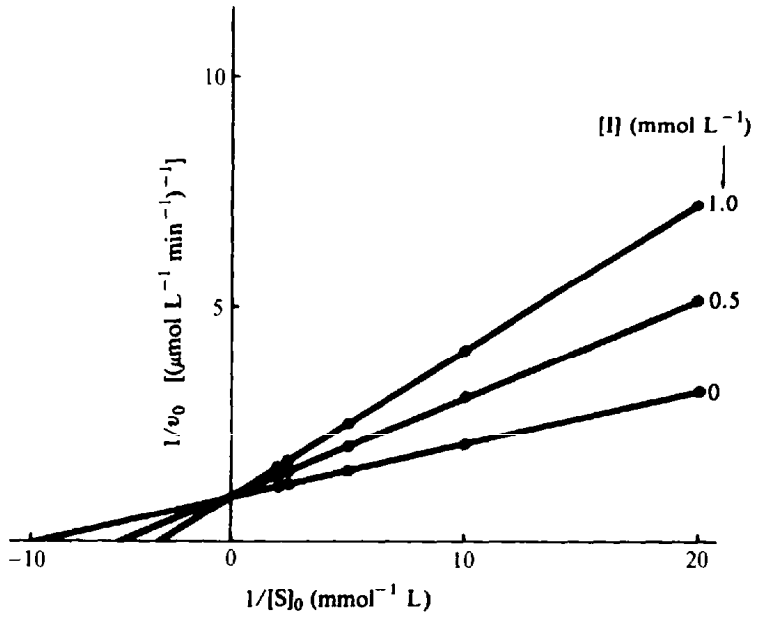


Fig. 9-13 Lineweaver-Burk plot of inhibitor effect.

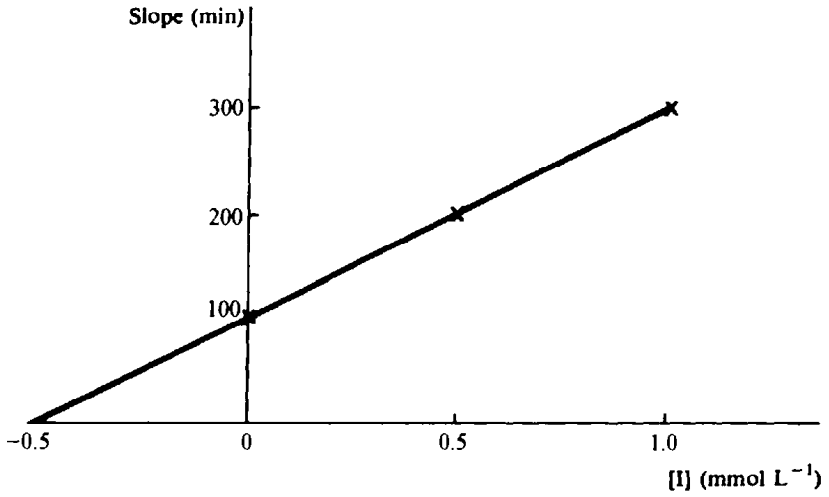


Fig. 9-14

intercept of this *secondary* plot on the abscissa is $-K_I$; hence K_I can be estimated as 0.5 mmol L^{-1}. The intercept on the ordinate is K_m/V_{max}. Since we know V_{max}, then K_m can be estimated as 0.1 mmol L^{-1}.

9.11. For each of the four types of enzyme inhibition given in Table 9.1, derive the Lineweaver-Burk equations and draw *archetypal* graphs.

SOLUTION

To derive the Lineweaver-Burk equations, we proceed by simply taking the reciprocals of each side of the equations expressing v_0 as a function of $[S]_0$ in Table 9-1. The corresponding graphs of $1/v_0$ versus $1/[S]_0$ have varying slopes, intercepts, or both as [I] is varied. Pure noncompetitive inhibition shows lines

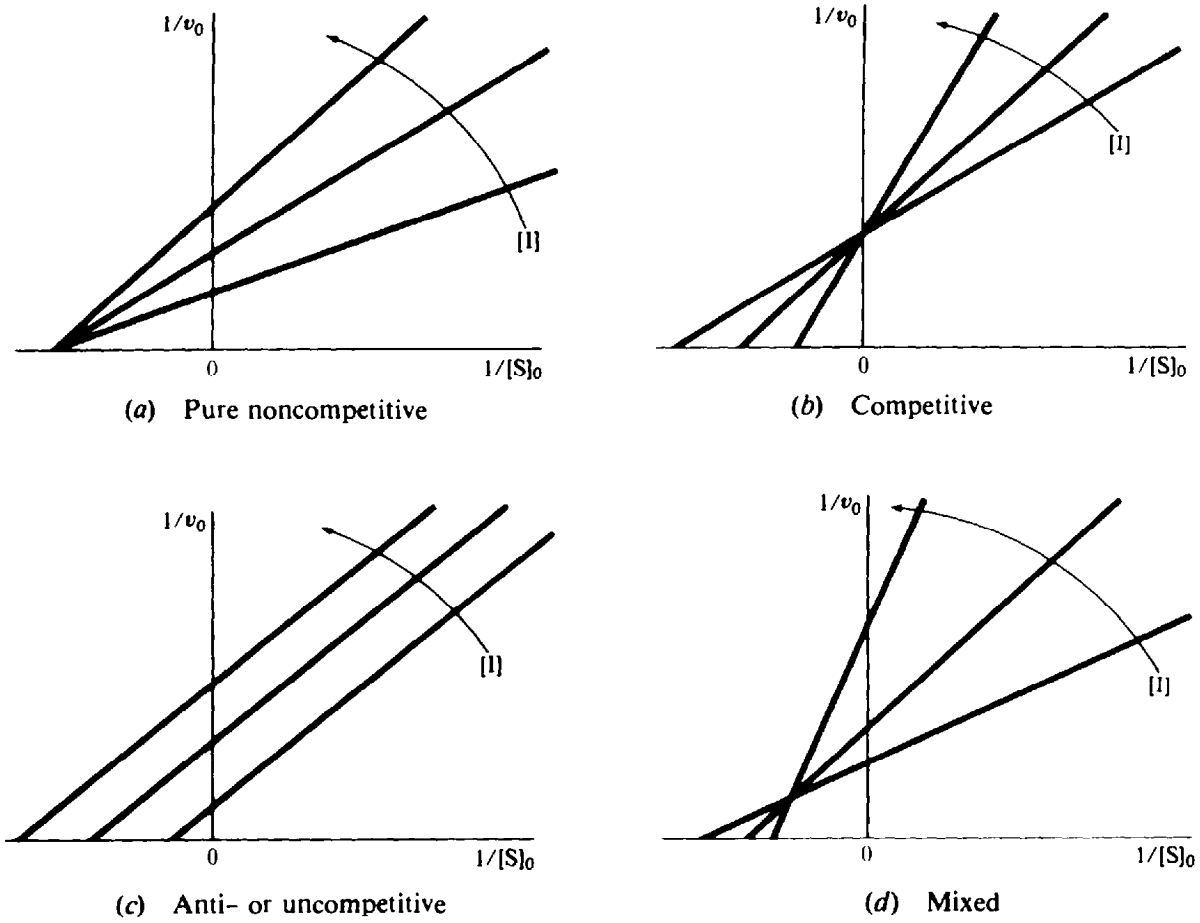


Fig. 9-15 Lineweaver-Burk plots for various types of enzyme inhibitors. The arrows indicate increasing [I].

intersecting on the abscissa. Competitive inhibition shows intersecting lines on the ordinate. For anti- or uncompetitive inhibition, the lines are parallel. For mixed inhibition, both slopes and axis intercepts are different for different values of [I]. See Fig. 9-15.

9.12. Consider the ordered Bi Bi reaction (Example 9.6):

$$A + B \rightleftharpoons P + Q$$

Predict the inhibition patterns that will be obtained when product P or Q is used as an inhibitor, with varied concentrations of A.

SOLUTION

We have four cases to consider. The effects of different inhibitor concentrations with:

(1) Varied [A]; fixed, nonsaturating [B]; P inhibitor

(2) Varied [A]; fixed, saturating [B]; P inhibitor

(3) Varied [A]; fixed, nonsaturating [B]; Q inhibitor

(4) Varied [A]; fixed, saturating [B]; Q inhibitor

It is convenient to consider the inhibition patterns in terms of plots of $1/v_0$ versus $1/[\text{substrate}]_0$ at various concentrations of inhibitor. Varying inhibitor may affect only the slopes of the plots (competitive inhibition),

both slopes and intercepts (noncompetitive or mixed inhibition), or only the intercepts (uncompetitive inhibition). Inhibition patterns can be predicted from the following so-called Cleland's rules:

(*a*) Competitive inhibition is obtained when the inhibitor and varied substrate bind to the *same* form of the enzyme, or different forms that are in *equilibrium* with each other.

(*b*) Noncompetitive or mixed inhibition is obtained when the inhibitor and varied substrate bind to *different* forms of the enzyme, linked by *reversible* interconversions of enzyme forms.

(*c*) Anti- or uncompetitive inhibition is obtained when the inhibitor and substrate bind to *different* forms of the enzyme, linked by *irreversible* interconversions of enzyme forms.

The reaction above can be written in full as:

$$E \underset{k_{-1}}{\overset{k_1 A}{\rightleftharpoons}} EA \underset{k_{-2}}{\overset{k_2 B}{\rightleftharpoons}} \begin{pmatrix} EAB \\ EPQ \end{pmatrix} \underset{k_{-3}P}{\overset{k_3}{\rightleftharpoons}} EQ \underset{k_{-4}Q}{\overset{k_4}{\rightleftharpoons}} E$$

Case 1 (varied [A], nonsaturating [B], P inhibitor): From rule (*b*) above, noncompetitive or mixed inhibition will be obtained, as A and P bind to different forms of the enzyme (E and EQ, respectively) linked by reversible interconversions.

Case 2 (varied [A], saturating [B], P inhibitor): In this case, the link between E and EQ is irreversible (because B is present at saturating levels); hence, from rule (*c*) above, uncompetitive inhibition will be obtained.

Case 3 (varied [A], nonsaturating [B], Q inhibitor): Both A and Q bind to the *same* form (E) of the enzyme; hence, from rule (*a*) above, competitive inhibition will be obtained.

Case 4 (varied [A], saturating [B], Q inhibitor): Although the interconversion of EA and EAB is irreversible in the direction of EAB formation at saturating [B], A and Q still bind to the *same* form of the enzyme (E); hence, from rule (*a*) above, competitive inhibition will be obtained.

Establishing the inhibition patterns in an enzyme-catalyzed reaction is usually an important step in elucidating the reaction mechanism. One complication in the interpretation of such data is the possible formation of *dead-end* complexes (i.e., a complex of the form EAP in the above scheme). This is especially important in *rapid-equilibrium* reactions [ones in which all steps except the rate constants for the central isomerization step (EAB EPQ in the above example) are very large].

pH EFFECTS ON ENZYME REACTION RATES

9.13. A hypothetical enzyme has activity, with respect to the hydrolysis of a neutral substrate, that is unaffected by pH over a very broad range, pH6 to pH8. However, it shows a much narrower pH-activity range when an alternative substrate, which contains an imidazole group, is used. A total of 50 assays were carried out by using 10 different substrate concentrations with a fixed amount of enzyme, at five different pH values. For each pH, the enzyme gave linear Lineweaver-Burk plots, and the estimates of K_m and V_{max} were as follows:

pH	K_m (mmol L^{-1})	V_{max} (μmol L^{-1} min^{-1}
8.0	1.1	75
7.5	1.3	80
7.0	2.0	77
6.5	4.2	83
6.0	11.0	75

(*a*) Describe the effect of varying pH on K_m and V_{max}. (*b*) What is the significance of the fact that the enzyme activity with an uncharged substrate scarcely alters with pH in the range 6–8? (*c*) Propose a reaction scheme that is consistent with the results. (*d*) Calculate the relevant dissociation constant of the H^+-reactant complex.

SOLUTION

(*a*) From the data it is evident that within experimental error V_{max} does not change with pH. However, K_m is strongly pH-dependent, with the enzyme having the highest affinity for the substrate at high pH.

(*b*) If the rate of an enzyme-catalyzed reaction is insensitive to pH, this suggests a lack of ionizable groups involved in binding, catalysis, or both. This is the case with the present enzyme. The fact that with the ionizable substrate the reaction rate is pH-dependent suggests that the ionization of the substrate affects the rate.

(*c*) In view of the above statements, a possible reaction scheme is

$$\begin{array}{ccccc} S + E & \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} & ES & \xrightarrow{k_2} & E + P \\ \Big\| K_s & & & & \\ SH^+ & & & & \end{array}$$

(*d*) The corresponding rate expression is

$$\frac{1}{v_0} = \frac{V_{max}[S]_0}{K_m\left(1 + \dfrac{[H^+]}{K_s}\right) + [S]_0}$$

This equation has the same form as Eq. (*9.35*) at low pH; i.e., Eq. (*9.37*). We see that V_{max} is independent of pH, but $K_m^{app} = K_m(1 + [H^+]/K_s)$. A plot of K_m^{app} versus $[H^+]$ for the data in the table yields K_m/K_s from the slope and K_m from the intercept; thus, $K_s = 10^{-7}$ mol L^{-1} and $K_m = 1.0$ mmol L^{-1}. See Fig. 9-16.

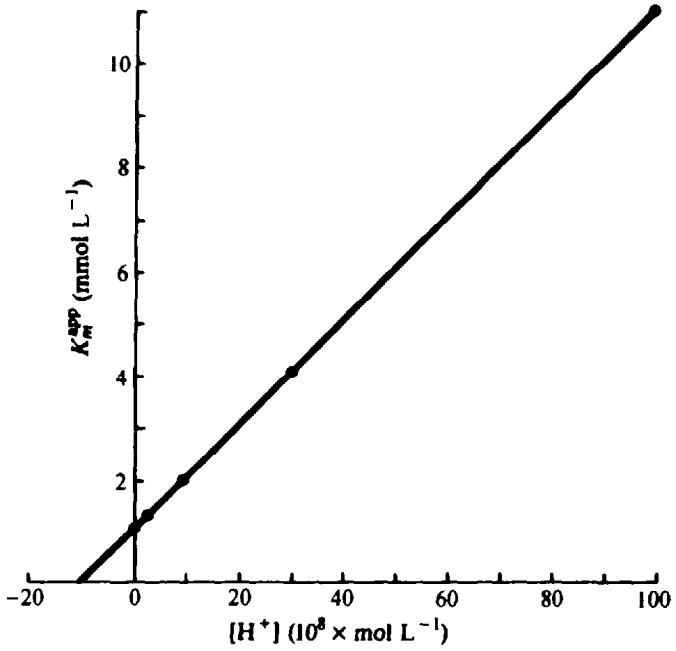


Fig. 9-16

9.14. Draw a reaction scheme that involves ionization of the *free enzyme*, rather than the substrate. Derive the corresponding rate equation that is similar to the one shown in Problem 9.13(*d*) above.

SOLUTION

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

$$K_e \Big\| \quad EH^+$$

The equation is

$$\frac{1}{v_0} = \frac{V_{max}[S]_0}{K_m\left(1 + \dfrac{[H^+]}{K_e}\right) + [S]_0}$$

This equation also has the same form as Eq. (*9.35*) at low pH [i.e., Eq. (*9.37*)], and is identical with the equation in Problem 9.13 above. Therefore it is impossible to distinguish between the present model (ionization of free enzyme) and that for ionization of the free substrate on the basis of the pH-dependence.

REGULATORY ENZYMES

9.15. The reaction scheme in Fig. 9-17 depicts isoleucine (E) synthesis from aspartate (A) by the bacterium *Rhodopseudomonas spheroides*. The control is called *sequential feedback control*. Describe the operation of this metabolic control system.

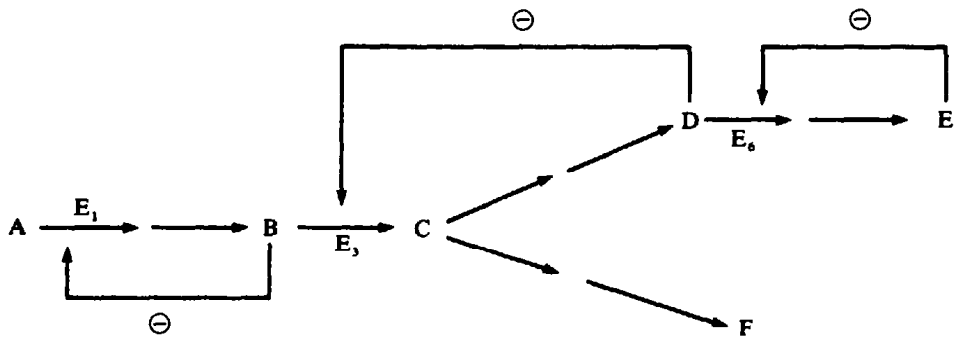


Fig. 9-17

SOLUTION

Overproduction of E (isoleucine) inhibits enzyme E_6 (threonine deaminase), and the consequent rise of D (threonine) reduces the rate of production of C (homoserine) via enzyme E_3 (homoserine dehydrogenase). The concentration of B (aspartate semialdehyde) rises, and this in turn inhibits E_1 (aspartokinase). It is therefore obvious why the control system is called a *negative feedback network*, or *sequential feedback system*.

9.16. In liver phosphofructokinase, ATP, ADP, and citrate are *effectors* of the reaction rate (see Fig. 9-18). Define what types of effectors they are.

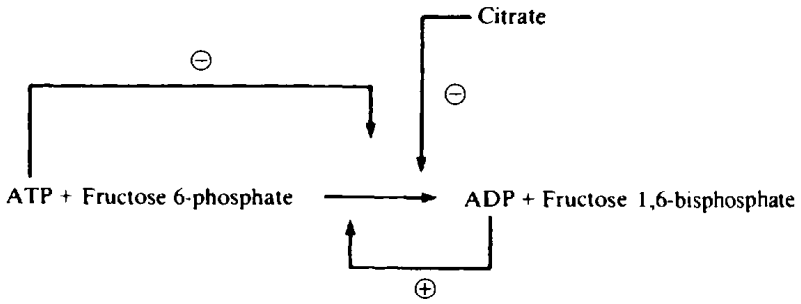


Fig. 9-18

SOLUTION

1. ATP: negative homotropic effector
2. Citrate: negative heterotropic effector
3. ADP: positive heterotropic effector

ATP exerts *negative feedforward* control (contrast with Example 9.7), while ADP exerts *positive feedback* control.

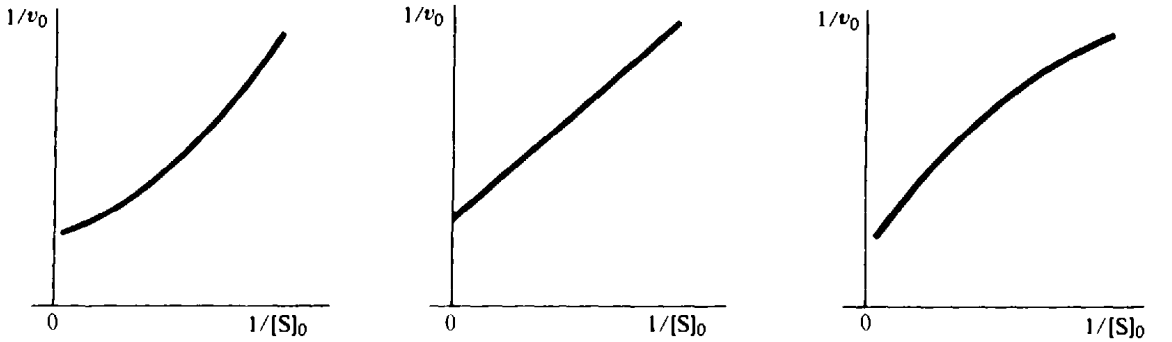
9.17. Describe the shapes of $1/v_0$ versus $1/[S]_0$ curves for enzymes exhibiting positive cooperativity, no cooperativity, and negative cooperativity of binding of the substrate (S).

SOLUTION

In plots of $1/v_0$ versus $1/[S]_0$, the slope of the line is directly proportional to K_m. In a positively cooperative case, the affinity of the enzyme for the substrate *increases*, and hence K_m *decreases* with increasing $[S]_0$ (or decreasing $1/[S]_0$). Therefore, the slope of the plot of $1/v_0$ versus $1/[S]_0$ will decrease as $1/[S]_0$ decreases [Fig. 9-19(*a*)].

In the case of *no cooperativity*, the affinity of the enzyme for the substrate, and hence K_m, is *constant* with changing $[S]_0$; therefore, the slope of the plot of $1/v_0$ versus $1/[S]_0$ will be constant [Fig. 9-19(*b*)].

In the case of *negative* cooperativity, the affinity of the enzyme for the substrate *decreases* with increasing $[S]_0$, hence K_m *increases* with increasing $[S]_0$. Therefore, the slope of the $1/v_0$ versus $1/[S]_0$ plot will *increase* with *decreasing* $1/[S]_0$ [Fig. 9-19(*c*)].



(*a*) Positively cooperative enzyme

(*b*) Noncooperative enzyme

(*c*) Negatively cooperative enzyme

Fig. 9-19

9.18. Prove that for a Michaelis-Menten enzyme, R_s of Eq. (*9.44*) is equal to 81.

SOLUTION

The rate equation is

$$v_0 = \frac{V_{max}[S]_0}{K_m + [S]_0}$$

and it is arranged to give an expression for $[S]_0$:

$$[S]_0 = \frac{v_0 K_m}{V_{max} - v_0}$$

$$R_s = \frac{0.9V_{max}K_m(V_{max} - 0.1V_{max})}{0.1V_{max}K_m(V_{max} - 0.9V_{max})}$$

$$= \frac{(0.9)(0.9)}{(0.1)(0.1)} = 81$$

9.19. If a solution contains two isozymes of an enzyme in equal amounts (as determined by gel electrophoresis) and it is known that the k_{cat} of each is the same but that the K_m for the substrate in one is 1 mmol L^{-1} and the other is 0.1 mmol L^{-1}, comment on the appearance of the Lineweaver-Burk plot of initial velocity data obtained for nine substrate concentrations ranging from 0.02 to 5 mmol L^{-1}.

SOLUTION

An *ideal* data set is generated from this expression:

$$v_0 = \frac{1[S]_0}{1 + [S]_0} + \frac{1[S]_0}{0.1 + [S]_0}$$

where $[S]_0 = [0.02, \ldots, 5]$ mmol L^{-1} $V_{max} = 1$ (arbitrary units)

$[S]_0$ (mmol L^{-1})	$1/[S]_0$ [(mmol L^{-1})$^{-1}$]	v_0 (arbitrary units)	$1/v_0$ [(arbitrary units)$^{-1}$]
0.020	50	0.19	5.3
0.025	40.0	0.22	4.5
0.04	25.0	0.32	3.1
0.10	10.0	0.59	1.7
0.20	5.0	0.83	1.2
0.50	2.0	1.17	0.85
1.00	1.0	1.41	0.71
2.00	0.5	1.63	0.61
5.0	0.2	1.82	0.55

The Lineweaver-Burk plot is shown in Fig. 9-20.

Notice that the line curves down from the right in Fig. 9-20. This is identical to the effect seen in negative cooperativity [see Fig. 9-19(*c*)]. In other words, this curvature is not diagnostic of *control* enzymes but can arise simply from a mixture of isoenzymes.

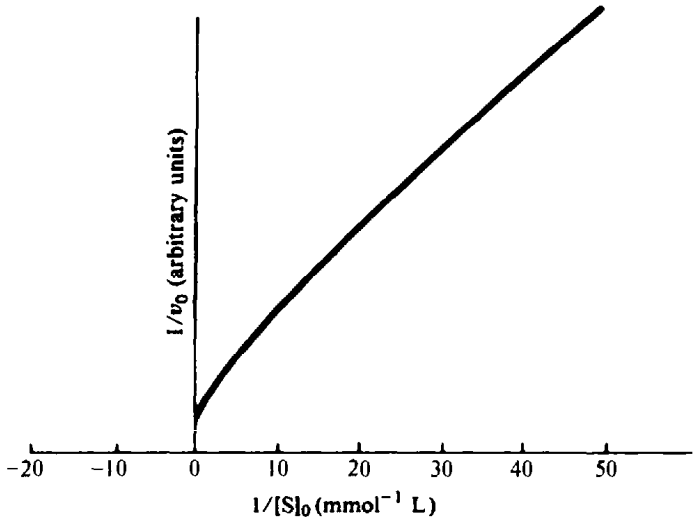


Fig. 9-20 Lineweaver-Burk plot for the sum of two rectangular hyperbolas: $V_{max1} = V_{max2} = 1$ and $K_{m1} = 1.0$ mmol L^{-1}, $K_{m2} = 0.1$ mmol L^{-1}.

9.20. Derive the Hill equation, Eq. (*9.45*).

SOLUTION

From the definition of K_b, Eq. (*9.46*), the concentration of the hemoglobin-O_2 complex is given by

$$[\text{Hb}(\text{O}_2)_n] = K_b[\text{Hb}](p\text{O}_2)^n$$

Y, the fractional saturation, is

$$Y = \frac{\text{concentration of Hb}(\text{O}_2)_n}{\text{total concentration of hemoglobin}}$$

$$= \frac{K_b[\text{Hb}](p\text{O}_2)^n}{[\text{Hb}] + K_b[\text{Hb}](p\text{O}_2)^n}$$

Cancellation of [Hb] from the numerator and denominator yields the Hill equation:

$$Y = \frac{K_b(p\text{O}_2)^n}{1 + K_b(p\text{O}_2)^n}$$

9.21. Prove that for $n > 1$, the Hill equation conforms to a positively cooperative binding or enzyme system.

SOLUTION

For an enzyme the Hill equation, rearranged, gives

$$[\text{S}]_0 = \left(\frac{v_0 K_m}{V_{max} - v_0}\right)^{1/n}$$

Therefore, from Prob. 9.18, $R_s = (81)^{1/n}$. So if $n > 1$, then $R_s < 81$, and according to the definition given in Sec. 9.11, under "Kinetic Behaviour of Regulatory Enzymes," the enzyme is positively cooperative.

9.22. The initial velocity of a reaction catalyzed by a regulatory enzyme was determined over the following range of initial substrate concentrations:

$[S]_0$ (mol L^{-1})	v_0 (mmol L^{-1} min^{-1})
0.1	1.6×10^{-4}
0.3	1.4×10^{-3}
0.5	4.0×10^{-3}
1.0	1.6×10^{-2}
3.0	0.14
5.0	0.39
10.0	1.45
100.0	14.6
500.0	15.9
1,000.0	16.0

Determine the Hill coefficient and apparent K_m.

SOLUTION

It is clear that V_{max} is ~16.0 mmol L^{-1} min^{-1} since the values of v_0 for $[S]_0$ of 500 and 1,000 mmol L^{-1} are very similar; thus, the enzyme is almost saturated at these high substrate concentrations. If we assume V_{max} has this value and use the expression in Example 9.11, we can construct the following table expressing $[S]_0$ as a molar concentration. A plot of these data gives Fig. 9-21 yielding a slope of n and an ordinate intercept of $\log_{10}(1/K_m)$; with these data $n = 2$ and $K_m = 1$ mmol L^{-1}.

$\log_{10}\left(\frac{v_0}{V_{max} - v_0}\right)$	$\log_{10}[S]_0$
−5.00	−4.00
−4.06	−3.52
−3.60	−3.30
−3.00	−3.00
−2.05	−2.52
−1.60	−2.30
−1.00	−2.00
+1.02	−1.00
+2.20	−0.30
—	0.00

9.23. What form does the MWC equation (*9.59*) take when the ratio of *intrinsic* binding constants, $K_T/K_R = c$, is equal to 1?

SOLUTION

The equation becomes

$$Y = \frac{\alpha(1+\alpha)^{n-1} + L\alpha(1+\alpha)^{n-1}}{(1+\alpha)^n + L(1+\alpha)^n}$$

By factoring out $(1+\alpha)^{n-1}$ in the numerator and $(1+\alpha)^n$ in the denominator, we obtain

$$Y = \frac{\alpha(1+L)(1+\alpha)^{n-1}}{(1+L)(1+\alpha)^n} = \frac{\alpha}{1+\alpha} = \frac{K_R[X]}{1+K_R[X]}.$$

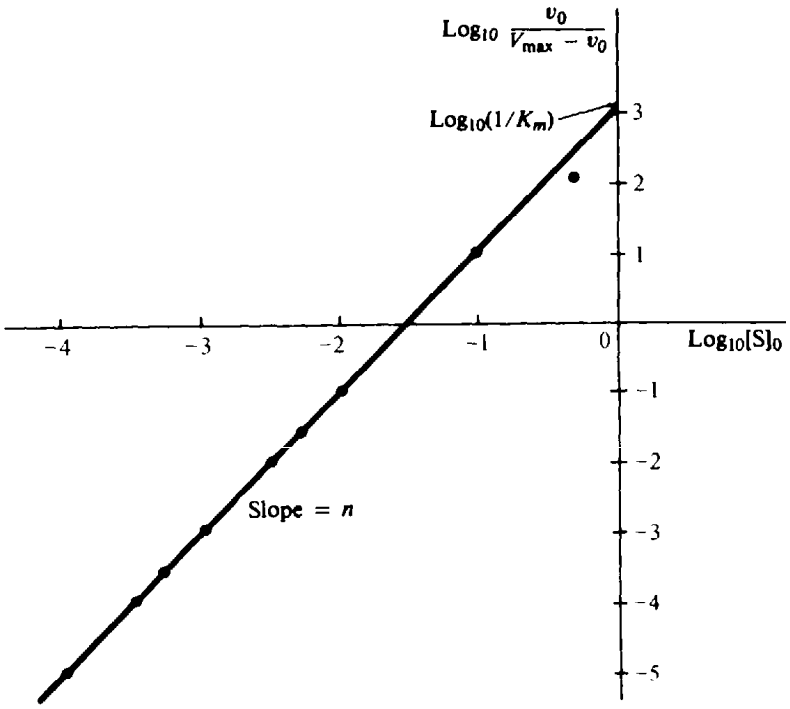


Fig. 9-21

Thus, this is a third situation in which this particular model reduces to a simple hyperbolic expression; the other two cases are discussed in Example 9.14. The outcome is physically "reasonable" since, if the ligand binds equally well to both states, the relative amounts of the states are irrelevant to the binding behavior.

Supplementary Problems

9.24. For the Michaelis-Menten equation there are algebraic transformations, in addition to the Lineweaver-Burk equation, that yield straight line plots from enzyme kinetic data. One such plot is due to Eadie and Hofstee; their equation takes the following form:

$$v_0 = -K_m \frac{v_0}{[S]_0} + V_{max}$$

Derive this equation.

9.25. An Eadie-Hofstee plot has values of v_0 on the ordinate and $v_0/[S]_0$ on the abscissa. Draw such a coordinate-axial system. How are K_m and V_{max} obtained from such a graph?

9.26. The Hanes-Woolf equation is another transformation of the Michaelis-Menten equation that yields a linear graph of the appropriate transformed variables. The equation is:

$$\frac{[S]_0}{v_0} = \frac{[S]_0}{V_{max}} + \frac{K_m}{V_{max}}$$

Derive this equation.

9.27. Construct a Hanes-Woolf plot and show how K_m and V_{max} are obtained from it.

9.28. (*a*) Draw the 5 percent coefficient of variation envelope around the data of Fig. 9-12. (*b*) Do this plot and analysis yield unbiased parameter estimates?

9.29. (*a*) Integrate the Michaelis-Menten equation and (*b*) derive a graphical procedure that enables the estimation of K_m and V_{max}.

9.30. (*a*) Integrate the noncompetitive inhibition Michaelis-Menten equation (Table 9.1) and (*b*) derive a graphical procedure that enables the estimation of K_I.

9.31. Prove that if an enzyme is inhibited competitively by its product and the numerical value of K_I is the same as that of K_m, the integrated equation from Table 9.1 may be written as

$$[S]_t = [S]_0 \exp[-V_{max} t/(K_m + [S]_0)]$$

which is a pure single exponential decay.

9.32. An ATPase from battong saliva was isolated, and the following ATP hydrolysis-rate data obtained, at an enzyme concentration of 10^{-8} mol L^{-1}:

[ATP] (μmol L^{-1})	5.0	1.7	1.0	0.7	0.56
v_0 (μmol L^{-1} min^{-1})	2.6	1.95	1.7	1.4	1.24

Determine K_m, V_m, and k_{cat} for the enzyme.

9.33. A *Theorell-Chance* enzyme mechanism is one in which the steady-state concentration of the central complex is effectively zero. In the case of a Bi Bi reaction, it may be represented as:

$$E \underset{k_{-1}}{\overset{k_1 A}{\rightleftharpoons}} EA \underset{k_{-2} P}{\overset{k_2 B}{\rightleftharpoons}} EQ \underset{k_3 Q}{\overset{k_3}{\rightleftharpoons}} E$$

Horse-liver *alcohol dehydrogenase* conforms to this basic scheme. Note that the central EAB/EPQ complex is not present. (*a*) Draw the reaction scheme by using the Cleland convention. (*b*) Show that the inhibition patterns for P and Q, with varied substrate A, both at saturating and at nonsaturating levels of B, are as follows:

P inhibition, nonsaturating B:	noncompetitive inhibition
P inhibition, saturating B:	no inhibition
Q inhibition, nonsaturating B:	competitive inhibition
Q inhibition, saturating B:	competitive inhibition

9.34. How does varying the substrate concentration relative to the enzyme concentration affect the validity of the steady-state assumption and the applicability of the Michaelis-Menten expression?

9.35. In a Michaelis-Menten enzyme mechanism, what substrate concentrations (relative to K_m) are needed for the reaction rate to be: (*a*) 0.1 V_{max}; (*b*) 0.25 V_{max}; (*c*) 0.5 V_{max}; (*d*) 0.9 V_{max}?

9.36. For an ionizable enzyme with a nonionizable substrate, prove that a plot of pK_m versus pH has a maximum at $pH = -0.5 \log(K_{1e} \cdot K_{2e})$.

9.37. Using Fig. 9-8, calculate R_S for (*a*) hemoglobin in red blood cells, (*b*) pure hemoglobin, and (*c*) pure myoglobin.

9.38. Can the value of n in the Hill equation, when it is fitted to real data, ever be less than 1?

9.39. Derive the two-site Adair equation.

Chapter 10

Metabolism: Underlying Theoretical Principles

10.1 INTRODUCTION

All living organisms transform energy taken from their surroundings. This energy is required for the synthesis of macromolecules to be used in growth and differentiation of the organism. These transformations are achieved via the action of a large number of enzymes (Chap. 8), catalyzing a complex network of chemical reactions collectively known as *metabolism*. This chapter reviews the theoretical principles underlying the study of this network of reactions; the detailed chemistry of metabolic processes is discussed in succeeding chapters.

10.2 THERMODYNAMICS

Living organisms can be considered as *physicochemical systems* interacting with their surroundings. *Thermodynamics* is the science of the energetics of such systems. It is a *macroscopic theory*, being concerned with the *bulk* properties of matter; the link between the thermodynamics and molecular processes is provided by the theory of *statistical mechanics*.

Basic Concepts

(*a*) *System*: that part of the universe with which we are concerned (e.g., an organism or a glass vessel in which a chemical reaction is occurring). The universe apart from the system is the system's *surroundings*.

(*b*) *Open system*: a system in which both matter and heat can exchange with the surroundings.

(*c*) *Closed system*: a system in which only *heat* can exchange with the surroundings.

(*d*) *Isolated (adiabatic) system*: a system in which neither matter nor heat can exchange with the surroundings.

EXAMPLE 10.1

Should living organisms be considered open, closed, or isolated systems?

Living organisms are *open* systems because they exchange both heat and matter (nutrients, excreta) with their surroundings.

(*e*) *State functions*: properties relating to changes in a system that are dependent *only* on its initial and final states. Many of the important properties of systems discussed below (e.g., internal energy, enthalpy, entropy, and Gibbs free energy) are state functions.

First Law of Thermodynamics

The first law of thermodynamics states that if a system exchanges heat with, or does work on, its surroundings, then a *change* in its *internal energy* (ΔU) occurs. This is expressed mathematically as

$$\Delta U = \Delta q - \Delta w \qquad (10.1)$$

where Δq is the heat exchanged with the surroundings and Δw is the work done on the surroundings.

EXAMPLE 10.2

What determines whether Δq and Δw are positive or negative?

If Δq is positive, heat has been transferred *to* the *system*, giving an *increase* in internal energy. When Δq is negative, heat has been transferred *to* the *surroundings*, giving a *decrease* in internal energy. When Δw is positive, work has been done by the system, giving a *decrease* in internal energy. When Δw is negative, work has been done by the surroundings, giving an *increase* in internal energy.

Question: How can ΔU values be measured experimentally?

For chemical reactions, ΔU values can be obtained by allowing the reaction to proceed in an *adiabatic calorimeter*, in which temperature changes during the reaction can be measured and converted to values for Δq. The most common source of work in chemical reactions is changes in volume (V) during the reaction. The work done against a pressure P is given by:

$$\Delta w = \int_{v_1}^{v_2} P\,dV \tag{10.2}$$

where v_1 and v_2 are the initial and final volumes of the system. Hence, from Eq. (*10.1*)

$$\Delta U = \Delta q - \int_{v_1}^{v_2} P\,dV \tag{10.3}$$

In a *constant volume* adiabatic calorimeter, dV is zero, and therefore $\Delta U = \Delta q_v$, while in a *constant pressure* adiabatic calorimeter, the above integral is $P\Delta V$, and hence:

$$\Delta U = \Delta q_p - P\Delta V \tag{10.4}$$

Question: How do calorimeters work?

In a *constant volume* adiabatic calorimeter (Fig. 10-1), the reaction of interest is initiated in a constant volume vessel, which is surrounded by a water bath. The calorimeter is well insulated, ensuring that there is no loss of heat from the system. The temperature change (ΔT) arising from

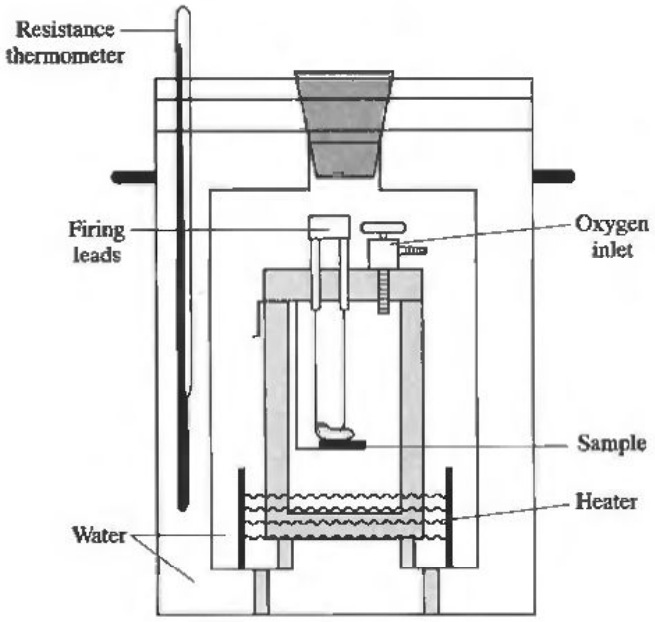


Fig. 10-1 Constant volume bomb calorimeter.

the completed reaction is converted to Δq_v using the *heat capacity* (C) of the calorimeter: $\Delta q_v = C\Delta T$. A *constant pressure* adiabatic calorimeter is based on the same principles, except the vessel in which the reaction occurs is designed to operate at constant pressure, rather than constant volume.

Enthalpy Changes

Many chemical reactions are conveniently studied at constant pressure, e.g., atmospheric pressure. From Eq. (*10.4*) above, the heat change during such a reaction is:

$$\Delta q_p = \Delta U + P\Delta V \tag{10.5}$$

This heat change is called the *enthalpy change* for a reaction and is symbolized by ΔH. Most biochemical reactions occur in solution; in these cases volume changes are negligible, and thus $\Delta H = \Delta U$.

EXAMPLE 10.3

The ΔU values for oxidation of glucose ($C_6H_{12}O_6$) and stearic acid ($C_{18}H_{36}O_2$) are -2.9×10^3 kJ mol^{-1} and -11.36×10^3 kJ mol^{-1}, respectively, at 310 K. (*a*) Calculate ΔH for these processes; (*b*) which of the two substances is more useful as an *energy store* in the body?

(*a*) For glucose the reaction is

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

In this case, there is no volume change (both O_2 and CO_2 are gases at 310 K). Hence, from Eq. (*10.5*), $\Delta U = \Delta H = -2.9 \times 10^3$ kJ mol^{-1}.

For stearic acid the reaction is

$$C_{18}H_{36}O_2 + 26O_2 \rightarrow 18CO_2 + 18H_2O$$

In this case, 8 moles of gas are consumed per mole of stearic acid ($26O_2 - 18CO_2$). Assuming that the ideal gas law applies, i.e.,

$$P\Delta V = \Delta n RT \tag{10.6}$$

where n is the number of moles of gas, and R is the gas constant (8.314 J mol^{-1} K^{-1}), and T is the absolute temperature, then

$$\Delta H = \Delta U + \Delta n RT = -\left(11.36 \times 10^3 + \frac{8 \times 8.31 \times 310}{1{,}000}\right) \text{ kJ mol}^{-1} = -11{,}381 \text{ kJ mol}^{-1}$$

(*b*) Clearly, a fat such as stearic acid is a far more energy-rich storage substance for a given number of carbon atoms than is glucose.

Second Law of Thermodynamics

The first law of thermodynamics provides a description of the energy balance for a given process; the second law provides a criterion for deciding whether or not the process will occur *spontaneously*. The second law of thermodynamics defines the *entropy change* (ΔS, in units of J K^{-1}) associated with a change in a closed system in terms of the heat absorbed by the system at constant temperature T;

$$\Delta S \geq \Delta q/T \tag{10.7}$$

The equality in this expression refers to a *reversible* process, while the inequality refers to an *irreversible* process.

Gibbs Free Energy

The second law of thermodynamics provides that a process will occur spontaneously if $\Delta q < T\Delta S$. For processes occurring at constant temperature and pressure, $\Delta q = \Delta H$ (see Eq. (*10.5*) above). Hence, for a spontaneous process

$$\Delta H - T\Delta S < 0 \tag{10.8}$$

The quantity $(\Delta H - T\Delta S)$ is symbolized by ΔG. Thus, if a process occurs spontaneously, $\Delta G < 0$. The symbol G refers to the *Gibbs free energy*, defined as

$$G = H - TS \tag{10.9}$$

and for a reaction at constant temperature and pressure

$$\Delta G = \Delta H - T\Delta S \tag{10.10}$$

The value of ΔG gives the maximum *work* available from the process, apart from work associated with changes in pressure or volume. ΔG may be expressed in joules or calories per mole (4.186 J = 1 cal).

EXAMPLE 10.4

Most of the energy transformed by higher animals derives from the oxidation of glucose:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

Given $\Delta H = -2{,}808\ \text{kJ mol}^{-1}$ and $\Delta S = 182.4\ \text{J K}^{-1}\ \text{mol}^{-1}$ for this reaction, how much energy is available from oxidation of 1 mole of glucose at 310 K?

We use:

$$\Delta G = \Delta H - T\Delta S = -(2{,}808 \times 10^3 + 310 \times 182.4)\ \text{J mol}^{-1} = -2{,}865\ \text{kJ mol}^{-1}$$

Hence, digestion of 1 mole (180.2 g) of glucose at 310 K provides an animal with 2,865 kJ of non-*PV* work.

Standard States

The *standard state* of a pure substance is defined as that form, at a specified temperature, that is stable at 1 atmosphere pressure (101.325 kPa). For *solutes*, the standard state is more conveniently defined as a 1 mol L^{-1} solution of the solute. For chemical reactions in solution, the *standard free energy change* ($\Delta G°$) is that for converting 1 mol L^{-1} of reactants into 1 mol L^{-1} of products:

$$\Delta G° = \Delta H° - T\Delta S° \tag{10.11}$$

Biochemical Standard States

Many biological processes involve hydrogen ions; the standard state of an H^+ solution is (by definition) a 1 mol L^{-1} solution, which would have a pH of nearly 0, a condition incompatible with most forms of life. Hence, it is convenient to define the *biochemical* standard state for solutes, in which all components except H^+ are at 1 mol L^{-1}, and H^+ is present at 10^{-7} mol L^{-1} (i.e., pH 7). Biochemical standard-state free energy changes are symbolized by $\Delta G°'$, and the other thermodynamic parameters are indicated analogously ($\Delta H°'$, $\Delta S°'$, etc.).

Equilibrium and Free Energy

The free energy change for the reaction

$$aA + bB \rightleftharpoons cC + dD$$

is given by

$$\Delta G = \Delta G^\circ + RT \ln\left(\frac{([C]/[C]^\circ)^c([D]/[D]^\circ)^d}{([A]/[A]^\circ)^a([B]/[B]^\circ)^b}\right) \tag{10.12}$$

where $[A]^\circ$, $[B]^\circ$, $[C]^\circ$, and $[D]^\circ$ are the concentrations of the reactants in their standard states. However, for reactions in solution, these concentrations are all $1\ \text{mol L}^{-1}$, hence Eq. (*10.12*) can be simplified to

$$\Delta G = \Delta G^\circ + RT \ln\left(\frac{[C]^c}{[A]^a}\frac{[D]^d}{[B]^b}\right) \tag{10.13}$$

If the above chemical reaction is at equilibrium (i.e., a reversible process), then $\Delta G = 0$. Hence, from Eq. (*10.13*)

$$\Delta G^\circ = -RT \ln\left(\frac{[C]_e^c}{[A]_e^a}\frac{[D]_e^d}{[B]_e^b}\right) \tag{10.14}$$

where the subscript e indicates the equilibrium concentrations of the reactants. However, the expression in parentheses is the *equilibrium constant* (K_e) for the reaction. Hence

$$\Delta G^\circ = -RT \ln K_e \tag{10.15}$$

EXAMPLE 10.5

The enzyme phosphoglucomutase catalyzes the reaction

Glucose 1-phosphate Glucose 6-phosphate

which has a ΔG° value of $-7.3\ \text{kJ mol}^{-1}$. If the enzyme is added to a $2 \times 10^{-4}\ \text{mol L}^{-1}$ solution of glucose 1-phosphate at 310 K, what will be the equilibrium composition of the solution?

From Eq. (*10.15*)

$$K_e = \exp(-\Delta G^\circ/RT) = \exp(7.3 \times 10^3/(8.31 \times 310)) = 17.00 \tag{10.16}$$

Hence

$$\frac{[\text{Glucose 6-phosphate}]_e}{[\text{Glucose 1-phosphate}]_e} = 17.00$$

and

$$[\text{Glucose 1-phosphate}] + [\text{Glucose 6-phosphate}] = 2 \times 10^{-4}\ \text{mol L}^{-1}$$

Therefore

$$[\text{Glucose 6-phosphate}]_e = 1.89 \times 10^{-4}\ \text{mol L}^{-1}$$
$$[\text{Glucose 1-phosphate}]_e = 1.11 \times 10^{-5}\ \text{mol L}^{-1}$$

Variation of Equilibrium Constant with Temperature

From Eq. (*10.15*), we have

$$\ln K_e = -\Delta G^\circ/RT \tag{10.17}$$

Hence

$$\frac{d \ln K_e}{dT} = -\frac{1}{R}\frac{d\left(\frac{\Delta G^\circ}{T}\right)}{dT} \tag{10.18}$$

Using Eq. (*10.11*)

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$$

we have

$$\frac{\Delta G^\circ}{T} = \frac{\Delta H^\circ}{T} - \Delta S^\circ \tag{10.19}$$

Assuming that $\Delta H°$ and $\Delta S°$ are *independent* of temperature

$$\frac{d\left(\frac{\Delta G°}{T}\right)}{dT} = -\frac{\Delta H°}{T^2} \tag{10.20}$$

Hence, from (*10.18*) and (*10.20*)

$$\frac{d\ln K_e}{dT} = \frac{\Delta H°}{RT^2} \tag{10.21}$$

This equation, the *van't Hoff isochore*, is often more useful in the integrated form:

$$\int_{K_1}^{K_2} d\ln K_e = \int_{T1}^{T_2} \frac{\Delta H°}{RT^2} dT \tag{10.22}$$

EXAMPLE 10.6

The enyme phosphorylase *b* binds AMP:

$$\text{Phosphorylase } b \cdot \text{AMP} \qquad \text{Phosphorylase } b + \text{AMP}$$

The equilibrium constant for this reaction is 2.75×10^{-5} mol L^{-1} at 286 K and 5.9×10^{-5} mol L^{-1} at 313 K. What are $\Delta H°$, $\Delta S°$, and $\Delta G°$ for the reaction at 303 K?

From Eq. (*10.22*), assuming that $\Delta H°$ is independent of temperature,

$$\ln\left(\frac{5.9 \times 10^{-5}}{2.75 \times 10^{-5}}\right) = \frac{-\Delta H°}{8.31}\left(\frac{1}{313} - \frac{1}{286}\right)$$

Hence $$\Delta H° = 21.03 \text{ kJ mol}^{-1}$$

Using this value, we can calculate the value of the equilibrium constant at 303 K

$$\ln\left(\frac{K_{303}}{5.9 \times 10^{-5}}\right) = \frac{-21.03 \times 10^3}{8.31}\left(\frac{1}{303} - \frac{1}{313}\right) = -0.2668$$

Therefore $$K_{303} = 4.5 \times 10^{-5} \text{ mol L}^{-1}$$

Now, substituting this value into the equation

$$\Delta G° = -RT\ln K_e$$

we have $$\Delta G°_{303\text{K}} = 25.2 \text{ kJ mol}^{-1}$$

Hence, using $$\Delta G° = \Delta H° - T\Delta S°$$

$\Delta S°$ can be calculated to be $-13.8 \text{ J K}^{-1} \text{ mol}^{-1}$.

10.3 REDOX REACTIONS

Chemical reactions involving oxidation and reduction processes (*redox* reactions) are central to metabolism. The energy derived from the oxidation of carbohydrates is coupled to the synthesis of ATP via a series of redox reactions, the mitochondrial electron-transport chain (see Chap. 14). Moreover, most life on earth is dependent on a series of redox reactions in *photosynthesis*, the process in which solar energy is used to produce ATP and O_2 and to synthesize carbohydrates from CO_2.

Basic Concepts

(*a*) Oxidation reaction: reaction in which a substance *loses* electrons.

(*b*) Reduction reaction: reaction in which a substance *gains* electrons.

(*c*) Half-cell reaction: the oxidation *or* reduction step in a *redox* reaction.

EXAMPLE 10.7

What are the half-cell reactions in the following redox reaction?

$$Zn + Cu^{2+} \rightleftharpoons Zn^{2+} + Cu$$

The *oxidation* half-cell reaction is

$$Zn \rightarrow Zn^{2+} + 2e^-$$

The *reduction* half-cell reaction is

$$Cu^{2+} + 2e^- \rightarrow Cu$$

Free Energy Changes in Redox Reactions

The free energy change for a redox reaction is given by

$$\Delta G = -nFE \tag{10.23}$$

where n is the number of electrons transferred from each molecule of the substance being oxidized to that being reduced, E is the *electromotive force* (in volts) required to prevent the electron transfer, and F is the *Faraday constant*, a conversion factor, equal to 96.5 kJ V^{-1} mol^{-1}. When the components of a redox reaction are in their standard states, Eq. (*10.24*) applies:

$$\Delta G^\circ = -nFE_0 \tag{10.24}$$

EXAMPLE 10.8

ΔG° for the reaction

$$Zn + Cu^{2+} \rightleftharpoons Zn^{2+} + Cu$$

is $-213\ kJ\ mol^{-1}$. What is E_0 for this reaction?

From Eq. (*10.24*), we have

$$E_0 = -\frac{\Delta G^\circ}{n\mathrm{F}} \tag{10.25}$$

$$= \frac{213}{2 \times 96.5} = 1.1\ V$$

(because *two* electrons are transferred).

Electrochemical Cells

Redox reactions can be studied using electrochemical cells. An electrochemical cell for the chemical reaction in Example 10.8 is shown in Fig. 10-2. The Cu and Zn electrodes dip into solutions of their respective ions and the *salt bridge* (containing concentrated KCl) maintains electrical contact between the two solutions. Electrons will flow from the Zn half-cell to the Cu half-cell if Zn is oxidized to Zn^{2+}, with concomitant reduction of Cu^{2+} to Cu in the Cu half-cell. The value of E for this reaction may be determined by measuring the potential difference (in volts) that has to be applied to the cell to prevent the electron flow.

Half-Cell Electrode Potentials

Half-cell reactions cannot be studied in isolation; all that can be measured is the difference in potential (ΔE) when two half-cells are linked to form an electrochemical cell. *Relative* electrode

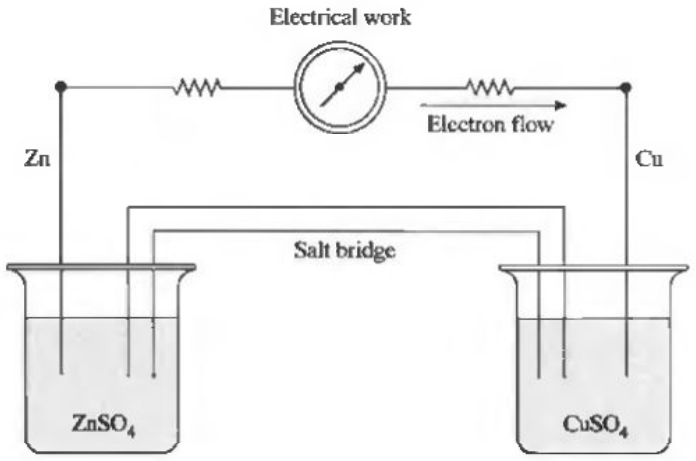


Fig. 10-2 Electrochemical cell.

potentials for half-cells are obtained by reference to a standard half-cell, the *hydrogen electrode*, which is assigned an E_0 of zero. The half-cell reaction for hydrogen is

$$H^+ + e^- \rightleftharpoons \tfrac{1}{2}H_2(Pt)$$

Hydrogen gas at 1 atm pressure is bubbled over a Pt electrode in a 1 mol L^{-1} H^+ solution. E_0 values for other half-cell reactions (with their components in their standard states) may be measured using electrochemical cells in which the hydrogen electrode is linked to an electrode at which the reaction of interest occurs.

By convention, the standard potentials of half-cells are expressed as *reduction potentials* (i.e., for the reduction reaction), such that larger negative values indicate a greater tendency to lose electrons.

The Relationship between E_0 and E

The free energy change for the general redox reaction occurring in solution

$$A_{ox} + B_{red} \rightarrow A_{red} + B_{ox}$$

is

$$\Delta G = \Delta G^\circ + RT \ln \left(\frac{[A_{red}][B_{ox}]}{[A_{red}]^\circ[B_{ox}]^\circ} \Big/ \frac{[A_{ox}][B_{red}]}{[A_{ox}]^\circ[B_{red}]^\circ} \right) \tag{10.26}$$

Using Eq. (*10.23*)

$$\Delta G = -nFE$$

and recalling that, for solutes, $[A_{red}]^\circ = [B_{ox}]^\circ = 1$ mol L^{-1}, etc. [cf. Eq. (*10.13*)], we have

$$E = E_0 - \frac{RT}{nF} \ln \left(\frac{[A_{red}][B_{ox}]}{B_{red}][A_{ox}]} \right) \tag{10.27}$$

This equation, the *Nernst redox equation*, provides a way of relating E and E_0 for any redox reaction or half-cell reaction.

EXAMPLE 10.9

What is E for the half-cell reaction

$$NAD^+ + H^+ + 2e^- \rightarrow NADH$$

at pH 7, 298 K, for 1 mol L^{-1} solutions of NAD^+ and NADH, given E° is -0.11 V?

The Nernst redox equation in this case is

$$E = E_0 - \frac{RT}{2F} \ln \left(\frac{[\text{NADH}]}{[\text{NAD}^+][\text{H}^+]} \right)$$

Substituting in values for E_0, [NADH], [NAD^+], etc., we get

$$E_0 = -0.11 - \frac{8.31 \times 298}{2 \times 9.65 \times 10^4} \ln(10^{+7}) = -0.32\ \text{V}$$

Since the components of the reaction are in their *biochemical* standard states (see Sec. 10.2), the above value for E is, in fact, defined as E_0' for this half-cell reaction (note the prime).

10.4 ATP AND ITS ROLE IN BIOENERGETICS

Properties of ATP

ATP is present in all living cells at concentrations of 10^{-3} to 10^{-2} mol L^{-1} of cell water. The ATP molecule is composed of three parts: adenine, D-ribose, and three phosphate groups in ester linkages (Chap. 7). The analogous compounds containing one and two phosphate groups are designated AMP and ADP, respectively.

ATP has a strong tendency to hydrolyze to ADP and phosphate; this is predicted from thermodynamics:

$$\text{ATP}^{4-} + \text{H}_2\text{O} \rightarrow \text{ADP}^{3-} + \text{HPO}_4^{2-} + \text{H}^+$$

Under biochemical standard-state conditions ($[\text{ATP}^{4-}] = [\text{ADP}^{3-}] = [\text{HPO}_4^{2-}] = 1$ mol L^{-1}, $[\text{H}^+] = 10^{-7}$ mol L^{-1}, and $T = 310$ K), the thermodynamic parameters for this reaction are $\Delta G^{\circ\prime} = -30.5$ kJ mol^{-1}, $\Delta H^{\circ\prime} = -20$ kJ mol^{-1}, and $\Delta S^{\circ\prime} = 34$ J K^{-1} mol^{-1}.

Question: Why is the hydrolysis of ATP thermodynamically favorable?

The *enthalpy* change for the reaction is favorable because (1) electrostatic repulsion between the negative charges in ATP exceeds that in the reaction products, (2) the reaction products are resonance stabilized, and (3) the enthalpies of solvation of the products are larger than that for ATP. The *entropy* change for the reaction is favorable because of the release of a phosphate group. Note that this implies that ATP hydrolysis is strongly temperature-dependent [cf. Eq. (*10.7*)].

ATP is often referred to as being *energy-rich* or containing *high-energy phosphate ester bonds*. This nomenclature should be thought of as implying that the overall process of ATP hydrolysis has large negative values of $\Delta G^{\circ\prime}$ and $\Delta H^{\circ\prime}$.

ATP in Bioenergetics

ATP is *not* a high-energy compound in comparison with many other biological compounds. The functions of ATP depend on its having a ΔG value for hydrolysis that is *intermediate* in value compared with ΔG values for hydrolysis of other phosphate esters. Thus, ATP and ADP can act as a *donor-acceptor* pair for *phosphoryl-group transfer*. In many cases the free energy of ATP hydrolysis is used to *support* reactions that would otherwise be thermodynamically unfavorable. This usually occurs via phosphorylation of one of the reactants in an otherwise *unfavorable* reaction.

EXAMPLE 10.10

The synthesis of glutamine from glutamate and NH_4^+ is thermodynamically unfavorable:

$$\text{Glutamate}^- + NH_4^+ \rightarrow \text{Glutamine} \qquad (\Delta G^{\circ\prime} = 14.2\ \text{kJ mol}^{-1})$$

How could the hydrolysis of ATP be used to render this reaction usable for the synthesis of glutamine?

The original reaction is altered by using ATP. The new reaction, which is thermodynamically favorable, consists of two *partial reactions* linked through a common intermediate via ATP hydrolysis:

$$\text{Glutamate}^- + ATP^{4-} \rightarrow \text{5-Phosphoglutamate}^{2-} + ADP^{3-}$$

$$\text{5-Phosphoglutamate}^{2-} + NH_4^+ \rightarrow \text{Glutamine} + \text{Phosphate}^-$$

The sum of these two reactions is:

$$\text{Glutamate}^- + ATP^{4-} + NH_4^+ \rightarrow \text{Glutamine} + ADP^{3-} + \text{Phosphate}^-$$

$\Delta G^{\circ\prime}$ for this reaction is $-16.3\ \text{kJ mol}^{-1}$, and hence it is thermodynamically favorable. The enzyme *glutamine synthetase* catalyzes this reaction in animal cells (Chap. 15).

10.5 CONTROL POINTS IN METABOLIC PATHWAYS

Enzymes that are subject to *control signals* generally fulfill two criteria: they are present at low enzymatic activities and catalyze reactions that are not at equilibrium under cellular conditions. Both criteria arise because control enzymes are likely to be those catalyzing the slowest (*rate-determining*) step in a metabolic pathway. This is likely to be the case if an enzyme is present at low activity. If this is the case, the enzyme-catalyzed reaction is unlikely to be at equilibrium in vivo because there is insufficient enzyme present to allow equilibration of its reactants before they react with other compounds.

Identification of Rate-Determining Steps in Metabolic Pathways

Identifying rate-determining steps can be done by measuring the activities of the enzymes in a metabolic pathway in vitro, using a homogenate of the tissue of interest and assay conditions [pH, temperature, ionic strength (Prob. 10.3), and substrate concentrations] similar to those in the cell. In some cases it is possible to measure enzyme activities in vivo using nuclear magnetic resonance spectroscopy (NMR).

Detection of Nonequilibrium Reactions

Nonequilibrium reactions can be detected by determining metabolite concentrations in the tissue of interest. Conventionally, a tissue sample is rapidly frozen by compression between metal plates that have been cooled to 77 K by immersion in liquid nitrogen (*freeze-clamping*). This procedure rapidly halts any enzymatic processes that might alter the metabolite concentrations; the concentrations can then be determined by enzymatic or chemical assays. Recently, ^{31}P-NMR spectroscopy has shown considerable value in measuring the concentrations of such metabolites as ATP, ADP, AMP, phosphate, and phosphocreatine in living cells or tissues.

Once the metabolite concentrations are known, the *mass-action ratio* (Γ) can be calculated. For the reaction

$$A + B + C + \cdots \rightleftharpoons P + Q + R + \cdots$$

$$\Gamma = \frac{[P][Q][R]\cdots}{[A][B][C]\cdots} \tag{10.28}$$

A comparison of Γ with the equilibrium constant for the reaction will establish whether or not the reaction is near equilibrium in vivo.

EXAMPLE 10.11

The following data were obtained for metabolite concentrations in a rat heart perfused with glucose:

Metabolite	Intracellular Concentration (mmol L^{-1})
ATP	11.5
ADP	1.3
AMP	0.17
Fructose 6-phosphate	0.09
Fructose 1,6-bisphosphate	0.02

Are the enzymes (*a*) phosphofructokinase and (*b*) adenylate kinase likely to be control enzymes in heart metabolism?

(*a*) For the phosphofructokinase reaction

$$\text{Fructose 6-phosphate} + \text{ATP} \xrightarrow{Mg^{2+}} \text{Fructose 1,6-bisphosphate} + \text{ADP}$$

$$\Gamma = \frac{[\text{fructose 1,6-bisphosphate}][\text{ADP}]}{[\text{fructose 6-phosphate}][\text{ATP}]}$$

$$= (0.02 \times 1.3)/(0.09 \times 11.5) = 0.025$$

This value is far smaller than that of the equilibrium constant for the reaction (~1,200); therefore, the reaction is *far* from equilibrium in rat heart, and hence phosphofructokinase is likely to be a control enzyme.

(*b*) For the adenylate kinase reaction

$$\text{ATP} + \text{AMP} \overset{Mg^{2+}}{\rightleftharpoons} 2\text{ADP}$$

$$\Gamma = \frac{[\text{ADP}]^2}{[\text{ATP}][\text{AMP}]} = (1.3)^2/(11.5 \times 0.17) = 0.86$$

This value is similar to that for the equilibrium constant for the reaction (~0.4); hence, this reaction is unlikely to be the control point of its metabolic pathway.

Note that we have assumed that all the intracellular metabolites are available to the enzymes mentioned; no account has been taken of possible effects of *metabolic compartmentation* (see Sec. 10.7). Hence, deductions about possible control enzymes, based on mass-action ratios, should be viewed with some caution.

The Crossover Theorem

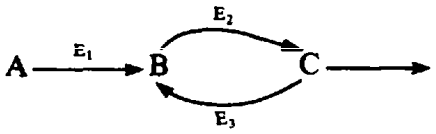
This principle relates to experiments in which a metabolic pathway is perturbed, for example, by adding an inhibitor or activator of one of the enzymes in the pathway. Following such a perturbation, the metabolite concentrations before and after a control enzyme in the pathway will change in opposite directions. This arises because such an enzyme catalyzes a slow, often rate-determining, step in the pathway. For example, if the perturbation *decreases* the activity of the enzyme, there will be an *increase* in the concentration of its substrate(s) and a *decrease* in the concentration of its product(s). This theorem is strictly valid only for nonbranching linear segments of metabolic pathways.

10.6 AMPLIFICATION OF CONTROL SIGNALS

The response of the *flux* (mass flow per unit time) through a metabolic pathway, but not the flux through a single enzyme-catalyzed reaction, to a regulatory signal can be subject to *amplification mechanisms*. Regulatory signals are usually changes in the concentrations of substrates or other effector molecules. These chemical signals can be amplified by *substrate cycles* or by cycles of *interconvertible enzymes*.

Substrate Cycles

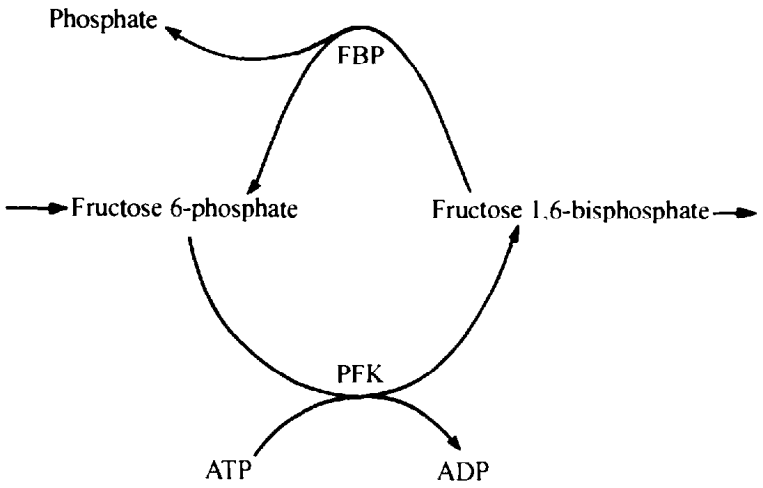
In the following segment of a metabolic pathway



a *substrate cycle* occurs between B and C: separate enzymes catalyze the interconversion of these two metabolites. The net flux through this step is the rate of E_2 minus the rate of E_3. If the rates of *both* enzymes are separately controlled, then the flux through the pathway is far more sensitive to changes in regulator concentration than if only one enzyme is involved.

EXAMPLE 10.12

In glycolysis, the enzymes phosphofructokinase (PFK) and fructose 1,6-bisphosphatase (FBP) form a substrate cycle:



The enzymatic activity of PFK is increased and that of FBP is decreased by elevated concentrations of AMP. The following data were obtained for the fractional saturation (Chap. 9) of the two enzymes with AMP:

AMP Concentration (arbitrary units)	Fractional Saturation
0	0
2.5	0.093
12.5	0.89

Calculate the change in flux through the pathway as [AMP] changes from 2.5 to 12.5 concentration units, assuming: that the maximum enzymatic activities for PFK and FBP are 100 and 10 mmol L^{-1} min^{-1}, respectively; that PFK is inactive in the absence of bound AMP; and that FBP is inactive when it has bound AMP.

The activity of PFK is given by

$$\text{PFK activity} = \text{maximum activity} \times \text{proportion of enzyme with bound AMP}$$

while that for FBP is given by

$$\text{FBP activity} = \text{maximum activity} \times \text{proportion of enzyme without bound AMP}$$

At an AMP level of 2.5 concentration units, we have

$$\text{PFK activity} = 100 \times 0.093 = 9.3 \text{ mmol } L^{-1} \text{ min}^{-1}$$

$$\text{FBP activity} = 10 \times (1.0 - 0.093) = 9.1 \text{ mmol } L^{-1} \text{ min}^{-1}$$

Hence, the relative flux through this section of glycolysis is

$$\text{Flux} = \text{PFK activity} - \text{FDP activity}$$

$$= 9.3 - 9.1 = 0.2 \text{ mmol } L^{-1} \text{ min}^{-1}$$

At an AMP level of 12.5 concentration units, we have analogously

$$\text{PFK activity} = 100 \times 0.89 = 89 \text{ mmol } L^{-1} \text{ min}^{-1}$$

$$\text{FBP activity} = 10 \times (1.0 - 0.89) = 1.1 \text{ mmol } L^{-1} \text{ min}^{-1}$$

$$\text{Flux} = 87.9 \text{ mmol } L^{-1} \text{ min}^{-1}$$

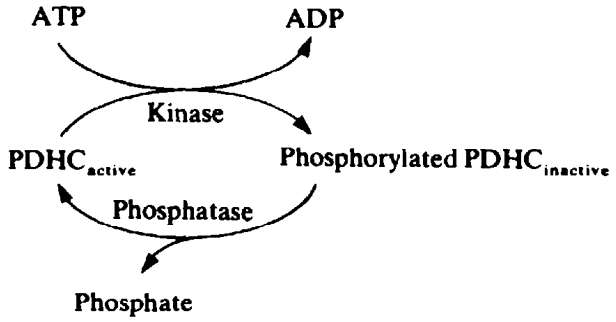
The substrate cycle has allowed a nearly 440-fold increase (87.9/0.2) in the flux through the pathway for only a 5-fold (12.5/2.5) change in the concentration of AMP.

Cycles of Interconvertible Enzymes

Some regulatory enzymes exist in different forms, the interconversions between these forms being enzyme-catalyzed. If the activities of the enzymes catalyzing the interconversions are controlled by changes in concentrations of effector molecules, then *amplification* of these chemical signals can occur.

EXAMPLE 10.13

The activity of the *pyruvate dehydrogenase complex* (PDHC) in mammals is altered by an enzyme-catalyzed phosphorylation (Chap. 12) of the complex



How could the activities of the kinase and phosphorylase be regulated so as to control the entry of pyruvate into the tricarboxylic acid cycle?

Pyruvate enters the tricarboxylic acid cycle after conversion into acetyl-CoA via the PDHC. The tricarboxylic acid cycle generates NADH from NAD^+, and the NADH then enters the mitochondrial electron-transport chain.

Hence, if either the ratios of acetyl-CoA/CoA or $NADH/NAD^+$ were high, it would be advantageous to render the PDHC inactive. There are three ways in which this is achieved: the kinase is active (1) when the $NADH/NAD^+$ ratio is high; (2) active when the acetyl-CoA/CoA ratio is high; and (3) the phosphatase is inhibited when the ratio of $NADH/NAD^+$ is high. Thus, the portion of the PDHC that is active is sensitive to the metabolic requirements of the cell.

10.7 INTRACELLULAR COMPARTMENTATION AND METABOLISM

Prokaryotic cells have volumes in the range 0.03–500 μm^3 and usually appear to contain no subcellular structures, whereas eukaryotic cells are larger, with volumes mostly in the range 50–15,000 μm^3, and contain many discrete subcellular organelles (Chap. 1). These organelles can be isolated, following disruption of the cells or tissue of interest, by differential centrifugation; organelles of different sizes and densities will sediment at different rates when placed in a centrifugal field.

Compartmentation of Metabolic Pathways

There are many interconnections between the main metabolic pathways. Many substrates and regulatory molecules, and some enzymes, are common to several pathways. An understanding of these interconnections requires knowledge of (1) the subcellular locations and concentrations of the enzymes involved, (2) the concentrations of metabolites within different subcellular organelles, and (3) the nature of permeability barriers for metabolites between the organelles; these barriers divide the cell into a number or compartments for each metabolite.

Location of Enzymes

The distribution of enzymes among the various subcellular organelles can be established by measurement of enzyme activities in the various fractions isolated by differential centrifugation. In this way, for example, the enzymes of glycolysis have been localized to the cytosol, while the distribution of the enzymes of the reverse process, gluconeogenesis (Chap. 11), has been found to be more complex. In this pathway, fructose bisphosphatase, together with the enzymes common to glycolysis, is cytosolic, while pyruvate carboxylase is found in the mitochondrial matrix, and phosphoenolpyruvate carboxykinase is cytosolic in rats, but is found in the mitochondria of pigeons.

Location of Metabolites

Establishing the concentrations of metabolites within different subcellular organelles or compartments is a complex task, because metabolite concentrations may change during the time required to isolate the organelles and low-molecular-weight metabolites may diffuse across organelle membranes during the isolation process. Most attention has been given to the concentrations of adenine and nicotinamide nucleotides and to intermediates of the tricarboxylic acid cycle in the cytosol and mitochondria. In some cases these concentrations can be estimated indirectly from measurements of enzyme-catalyzed reactions that are at equilibrium.

EXAMPLE 10.14

Lactate dehydrogenase is a cytosolic enzyme that is present at high enough concentrations for its reaction to be at equilibrium in many cells.

$$\text{Lactate} + NAD^+ \rightleftharpoons \text{Pyruvate} + NADH + H^+$$

The equilibrium constant for this reaction is known from in vitro studies. How could this be used to calculate the cytosolic $NADH/NAD^+$ ratio?

The cytosolic concentrations of pyruvate and lactate can be obtained from the results of freeze-clamping experiments. Hence, from the known $[H^+]$ of the cytosol and the known equilibrium constant (K_e)

$$K_e = \frac{[\text{Pyruvate}][\text{NADH}][H^+]}{[\text{Lactate}][\text{NAD}^+]}$$

the $NADH/NAD^+$ ratio can be estimated as ~1:1,000.

Solved Problems

THERMODYNAMICS

10.1. A sample of acetic acid (1 mol) was completely oxidized to CO_2 and H_2O in a constant volume adiabatic calorimeter, at 298 K. The heat released in the oxidation was 874 kJ. Calculate ΔH for the oxidation of acetic acid.

SOLUTION

At constant volume, the heat absorbed by the system (Δq_v) is equal to ΔU. Hence, because heat is released (transferred *from* the system *to* the surroundings)

$$\Delta U = -\Delta q_v = -874 \text{ kJ mol}^{-1}$$

ΔU is related to ΔH by the equation

$$\Delta H = \Delta U + P\Delta V$$

However, given the reaction for oxidation of acetic acid

$$CH_3COOH + 2O_2 \rightarrow 2CO_2 + 2H_2O$$

there is no volume change (as both O_2 and CO_2 are gases at 298 K), and hence ΔH for the oxidation is -874 kJ mol^{-1}.

10.2. The ΔH° value for the complete oxidation of fumaric acid (HOOC—CH=CH—COOH) is $-1{,}336 \text{ kJ mol}^{-1}$. Given the following data:

$$C + O_2 \rightarrow CO_2 \qquad \Delta H^\circ = -393 \text{ kJ mol}^{-1}$$

$$H_2 + \tfrac{1}{2}O_2 \rightarrow H_2O \qquad \Delta H^\circ = -285.5 \text{ kJ mol}^{-1}$$

calculate the enthalpy of formation (ΔH°) of fumaric acid from its elements.

SOLUTION

Because enthalpy is a state function, we can calculate the ΔH° of fumaric acid formation by summation of the enthalpy changes for the following reactions:

	ΔH°
$4C + 4O_2 \rightarrow 4CO_2$	$-1{,}572$ kJ
$2H_2 + O_2 \rightarrow 2H_2O$	-571 kJ
$4CO_2 + 2H_2O \rightarrow C_4H_4O_4 + 3O_2$	$+1{,}336$ kJ

Net equation:

$$4C + 2H_2 + 2O_2 \rightarrow C_4H_4O_4 \qquad \Delta H^\circ = -807 \text{ kJ mol}^{-1}$$

This procedure, which is an example of *Hess's law of constant heat summation*, relies upon the fact that ΔH for any reaction depends only on the final and initial states, and illustrates a convenient method for calculating ΔH values that would be difficult to measure experimentally.

10.3. The respective concentrations of Na^+ and K^+ ions in tissues are approximately 10 mmol L^{-1} and 90 mmol L^{-1} inside cells and 140 mmol L^{-1} and 4 mmol L^{-1} outside cells. Calculate the free-energy requirements for maintenance of these ion gradients.

SOLUTION

We can represent the ion gradients by the equilibria:

$$Na^+_{in} \qquad Na^+_{out}$$

$$K^+_{in} \qquad K^+_{out}$$

In both cases, the ΔG values are given by the expression:

$$\Delta G = \Delta G^\circ + RT \ln\left(\frac{[\text{ion}]_{out}}{[\text{ion}]_{in}}\right)$$

$$\Delta G^\circ = -RT \ln\left(\frac{[\text{ion}]_{out_e}}{[\text{ion}]_{in_e}}\right)$$

where the subscript e indicates the equilibrium concentrations of the ions. However, at equilibrium one would expect equal concentrations of the ions inside and outside the cells. Hence, $\Delta G^\circ = 0$, and therefore the free energy required to maintain the ion gradients is:

$$\Delta G = RT \ln\left(\frac{[\text{ion}]_{out}}{[\text{ion}]_{in}}\right)$$

Substituting the Na^+ and K^+ concentrations into this equation, we have

$$\Delta G_{Na^+} = 6.8 \text{ kJ mol}^{-1}$$

$$\Delta G_{K^+} = 8.0 \text{ kJ mol}^{-1}$$

We have made a number of assumptions in this calculation, the most notable being that the ionic solutions are *ideal*, in that there are no interactions (attractive or repulsive) between solute molecules. It is most unlikely that this is the case, especially in moderately concentrated solutions of ions. In order to correct for nonideality (interactions between solute molecules), we need to substitute the *activities* of solute molecules for their *concentrations* in all thermodynamic calculations. The activity (a) of a solute molecule is related to its concentration (C) by an *activity coefficient* (γ).

$$a = \gamma C \tag{10.29}$$

For solutions of electrolytes, γ can be calculated using *Debye-Hückel theory*. The mean activity coefficient ($\gamma_\pm$) of positively and negatively charged ions in solution at 25°C is given by:

$$\log \gamma_\pm = -0.5 z_+ z_- I^{1/2} \tag{10.30}$$

where z_+ and z_- are the charges carried by the ions and I is the *ionic strength* of the solution:

$$I = \tfrac{1}{2} \Sigma C_i z_i^2 \tag{10.31}$$

where the sum is carried out over all ion types (i) in the solution.

10.4. Calculate the mean activity coefficient in a solution of 0.25 mol L^{-1} Na_3PO_4.

SOLUTION

This salt is completely dissociated in water. First, we calculate the ionic strength from Eq. (10.31):

$$I = \tfrac{1}{2}(C_{Na^+} z_{Na^+}^2 + C_{PO_4^{3-}} z_{PO_4^{3-}}^2) = \tfrac{1}{2}(0.75 \times 1^2 + 0.25 \times 3^2) = 1.5 \text{ mol L}^{-1}$$

Now, using Eq. (*10.30*) with $z_+ = 1$, and $z_- = 3$:

$$\log \gamma_{\pm} = -1.837$$

$$\gamma_{\pm} = 0.0146$$

Strictly, this value is valid only at 298 K, as the constant (0.5) in the equation for $\log \gamma_{\pm}$ is (slightly) different at other temperatures.

10.5. Consider the general dissociation of an acid HA

$$HA \rightleftharpoons H^+ + A^-$$

Derive an expression for the dependence of the concentrations of dissociated and undissociated acid ($[A^-]$ and [HA]) on pH.

SOLUTION

The equilibrium constant for the dissociation is

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

Hence,

$$\log K_a = \log[H^+] + \log\left(\frac{[A^-]}{[HA]}\right)$$

and since $pH = -\log[H^+]$

$$-\log K_a = pH - \log\left(\frac{[A^-]}{[HA]}\right)$$

Substituting pK_a for $-\log K_a$

$$pK_a = pH - \log\left(\frac{[A^-]}{[HA]}\right)$$

This is the Henderson-Hasselbalch equation (Chap. 3).

REDOX REACTIONS

10.6. Calculate $\Delta G^{\circ\prime}$ for the following reaction:

$$\text{Pyruvate} + \text{NADH} + H^+ \qquad \text{Lactate} + \text{NAD}^+$$

The half-cell reactions are:

$$\text{Pyruvate} + 2H^+ + 2e^- \rightarrow \text{Lactate} \qquad E_0' = -0.19\ \text{V}$$

$$\text{NAD}^+ + H^+ + 2e^- \rightarrow \text{NADH} \qquad E_0' = -0.32\ \text{V}$$

SOLUTION

From the half-cell reactions, E_0' for the reaction is $-0.19 - (-0.32) = 0.13$ V. Now, using the equation

$$\Delta G^{\circ\prime} = -nFE_0'$$

with $n = 2$ in this case,

$$\Delta G^{\circ\prime} = -2 \times 9.65 \times 10^4 \times 0.13\ \text{J mol}^{-1} = -25.1\ \text{kJ mol}^{-1}$$

Note that a *positive* E_0' value for a redox reaction implies that it is thermodynamically *favorable*, having a *negative* value of $\Delta G^{\circ\prime}$.

10.7. The *electrochemical potential* μ of an ion of charge z in an *electrostatic field* ψ is (assuming ideal behavior) defined by the equation

$$\mu = zF\psi + \mu^\circ + RT \ln X \tag{10.32}$$

where μ° is the *chemical potential* (free energy per mole) of the ion in its standard state and X is the *mole fraction* of the ion, given by the expression

$$X = \frac{\text{(moles of ion)}}{\text{(moles of ion)} + \text{(moles of solvent)}} \tag{10.33}$$

For dilute solutions (moles of solvent) $\gg$ (moles of ion); hence

$$X = \frac{\text{moles of ion}}{\text{moles of solvent}} \tag{10.34}$$

In the *chemiosmotic theory* for oxidative phosphorylation (Chap. 14), electron flow in the electron-transport chain is coupled to the generation of a proton concentration gradient across the inner mitochondrial membrane. Derive an expression for the difference in electrochemical potential for a proton across the membrane.

SOLUTION

The difference in chemical potential is given by the expression

$$\Delta\mu = \mu_{in} - \mu_{out}$$

From Eq. (*10.32*),

$$\Delta\mu = zF(\psi_{in} - \psi_{out}) + RT \ln\left(\frac{[H^+]_{in}}{[H^+]_{out}}\right)$$

Hence

$$\Delta\mu = zF\Delta\psi - 2.3RT\Delta\text{pH} \tag{10.35}$$

where $\Delta\psi$ is the potential difference across the membrane, and ΔpH the pH difference across the membrane. This is the *Nernst potential equation*.

ATP AND ITS ROLE IN BIOENERGETICS

10.8. The negatively charged phosphate groups of ATP provide an effective means of chelating divalent cations such as Mg^{2+}. Given that the following reaction

$$Mg^{2+} + ATP^{4-} \rightleftharpoons MgATP^{2-}$$

has an association equilibrium constant of $10^4\,L\,mol^{-1}$, calculate the proportion of ATP in a cell that is present as the Mg^{2+} chelate if the free Mg^{2+} concentration is $2 \times 10^{-2}\,mol\,L^{-1}$.

SOLUTION

$$\frac{[MgATP^{2-}]}{[Mg^{2+}][ATP^{4-}]} = 10^4\,L\,mol^{-1}$$

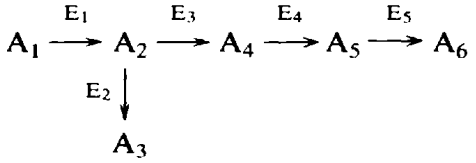
Hence,

$$\frac{[MgATP^{2-}]}{[ATP^{4-}]} = 2 \times 10^2$$

Therefore, more than 99 percent of the ATP in a cell is present as the Mg^{2+} chelate. It is thus not surprising that most ATP-utilizing enzymes use $MgATP^{2-}$, rather than ATP^{4-}, as a substrate.

CONTROL POINTS IN METABOLIC PATHWAYS

10.9. Given the following hypothetical scheme for the enzyme-catalyzed interconversions of metabolites A_1–A_6:



which enzyme in the scheme would be most suitable for being subject to metabolic control in terms of regulating the levels of the metabolite A_6?

SOLUTION

If enzyme E_1 were subject to feedback inhibition at high levels of metabolite A_6, then high levels of A_6 would lead to decreased production of metabolite A_2. This could have the undesirable effect of decreasing the production of metabolite A_3, via the action of enzyme E_2. This problem is overcome if enzyme E_3, which catalyzes the *first committed step* in the production of A_6, is subject to metabolic control.

AMPLIFICATION OF CONTROL SIGNALS

10.10. Does the existence of the phosphofructokinase–fructose-1,6-bisphosphatase *substrate cycle* have any energetic disadvantages?

SOLUTION

This substrate cycle has the disadvantage of acting, in effect, as an *ATPase*; ATP is converted to ADP and phosphate without net production of fructose-1,6-bisphosphate. This may be outweighed by the advantage conferred on an animal of having an efficient control mechanism in the form of a substrate cycle.

INTRACELLULAR COMPARTMENTATION AND METABOLISM

10.11. The enzyme succinate dehydrogenase (Chaps. 12 and 14) catalyzes the reaction:

$$\text{Succinate} + \text{FAD} \rightarrow \text{Fumarate} + \text{FADH}_2$$

A sample of rat liver was homogenized, the various cell organelles were isolated by differential centrifugation, and the succinate dehydrogenase enzymatic activity in each organelle preparation was determined:

Organelle Preparation	Succinate Dehydrogenase Activity (% of total)
Nuclei	11
Mitochondria	75
Lysosomes	6
Plasma membrane	5
Cytosol	3

How could these data be useful in investigating intracellular compartmentation?

SOLUTION

The data show that succinate dehydrogenase is found primarily in the mitochondria; it can thus act as a *marker enzyme* for these organelles. Marker enzymes can be used to determine the extent of contamination of a particular subcellular fraction by other organelles; in this case, for example, it would appear that the preparation of nuclei is contaminated with mitochondria.

Supplementary Problems

10.12. The *osmotic pressure* (π) of a solution is given by the equation

$$\pi = RTa \qquad (10.36)$$

where a is the activity of the solute. Sketch the curve you would expect to see for the relationship between the osmotic pressure of a solution of NaCl and the concentration of NaCl.

10.13. Given the half-cell potentials

$$\tfrac{1}{2}O_2 + 2H^+ + 2e^- \rightleftharpoons H_2O \qquad E_0' = 0.82\,V$$

$$NAD^+ + H^+ + 2e^- \rightleftharpoons NADH \qquad E_0' = -0.32\,V$$

calculate $\Delta G^{\circ\prime}$ across the electron-transport chain.

10.14. Given that $\Delta G^{\circ\prime}$ for the synthesis of ATP from ADP and phosphate is 30.5 kJ mol^{-1}, and assuming that two protons are translocated by the mitochondrial ATPase per molecule of ATP synthesized, use Eq. (*10.35*) in Prob. 10.7 to calculate the *minimum* value of $\Delta\mu$ necessary for synthesis of ATP. (*Hint*: $\Delta\mu$ has units of kJ mol^{-1}; how can this be converted to units of mV to give $\Delta\mu$ as proton-motive force, Δp?)

10.15. Calculate the ionic strengths of solutions of Na_2SO_4 at (*a*) 0.01 mol L^{-1}; (*b*) 0.05 mol L^{-1}; (*c*) 0.1 mol L^{-1}; and (*d*) 1.0 mol L^{-1}.

10.16. The enzyme pyruvate carboxylase (Chap. 12) from chicken liver is an oligomer composed of four identical subunits. The enzyme loses its catalytic activity when cooled below 277 K. Assuming that this loss of activity reflects dissociation of the tetrameric enzyme into its subunits, what can you deduce about the relative importance of enthalpic and entropic effects in the association of the subunits in the tetrameric enzyme?

10.17. (*a*) For the following alcohol dehydrogenase-catalyzed reaction

$$CH_3CH_2OH + NAD^+ \rightleftharpoons CH_3CHO + NADH + H^+$$

derive an expression for the variation of ΔG° with pH. (*b*) Given that $\Delta G^{\circ\prime}$ for this reaction is 18.5 kJ mol^{-1}, what is ΔG° at a pH of 6 ($T = 298$ K)?

10.18. For the ionization of acetic acid

$$CH_3COOH \rightleftharpoons CH_3COO^- + H^+$$

ΔG° and ΔH° values at 298 K are 27.1 kJ mol^{-1} and -0.39 kJ mol^{-1}. (*a*) Calculate ΔS° for the reaction at 298 K. (*b*) Evaluate the equilibrium constant for the reaction at 323 K, noting any assumptions necessary for the calculation.

10.19. The hydrolysis of acetyl phosphate

$$\text{Acetyl phosphate} \rightleftharpoons \text{Acetate} + \text{Phosphate}$$

has a ΔG° value of $-42.3\,\text{kJ mol}^{-1}$ at 298 K. In principle, which of the following phosphate ester hydrolysis reactions could be used to synthesize acetyl phosphate by a mechanism analogous to that in Example 10.10?

(*a*) 1,3-Bisphosphoglycerate → 3-Phosphoglycerate + Phosphate
($\Delta G^{\circ\prime} = -57\,\text{kJ mol}^{-1}$)

(*b*) Creatine phosphate → Creatine + Phosphate
($\Delta G^{\circ\prime} = -42.7\,\text{kJ mol}^{-1}$)

(*c*) ADP → AMP + Phosphate
($\Delta G^{\circ\prime} = -27.6\,\text{kJ mol}^{-1}$)

10.20. The *control strength* (C) provides a way of describing relative changes in the flux through a metabolic pathway arising from changes in the activity of an enzyme in the pathway. The control strength is defined as:

$$C_i = \frac{\partial \ln \nu_g}{\partial \ln \nu_i} = \frac{\partial \nu_g}{\delta \nu_i} \cdot \frac{\nu_i}{\nu_g} \tag{10.37}$$

where ν_g is the net flux through the pathway, and ν_i that through the step catalyzed by enzyme E_i. If the substrates of the enzyme are present in large excess, show that

$$C_i = \frac{[E_i]}{\nu_g} \cdot \frac{\partial \nu_g}{\partial [E_i]} \tag{10.38}$$

(*Hint*: The control strength can also be written as

$$C_i = \nu_i(\partial \nu_g/\partial p)/\nu_g(\partial \nu_i/\partial p)$$

where p is an arbitrary parameter. What is the relationship between ν_i and $[E_i]$ at high substrate concentrations?)

10.21. Prove that the sum of the control strengths of all enzymes in a linear sequence of reactions is 1.

Chapter 11

Carbohydrate Metabolism

11.1 GLYCOLYSIS

Glycolysis is a process that results in the conversion of a molecule of glucose into two molecules of pyruvate. It is a primitive metabolic pathway since it operates in even the simplest cells and does not require oxygen. The pathway of glycolysis performs *five* main functions in the cell:

1. Glucose is converted to pyruvate, which can be oxidized in the citric acid cycle (Chap. 12).
2. Many compounds other than glucose can enter the pathway at intermediate stages.
3. In some cells the pathway is modified to enable glucose to be synthesized.
4. The pathway contains intermediates that are involved in alternative metabolic reactions.
5. For each molecule of glucose that is consumed, two molecules of ADP are phosphorylated by *substrate-level* phosphorylation to produce two molecules of ATP.

Question: What is the overall, balanced chemical equation for glycolysis?

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + 2ADP + 2NAD^+ + 2P_i \rightarrow \underset{\text{Pyruvate}}{2C_3H_4O_3} + 2ATP + 2NADH + 2H^+ + 2H_2O$$

The apparent simplicity of this equation conceals the complexity of the glycolytic pathway, which involves nine intermediate compounds and 10 enzymes; the enzymes are located in the cytoplasm of the cell.

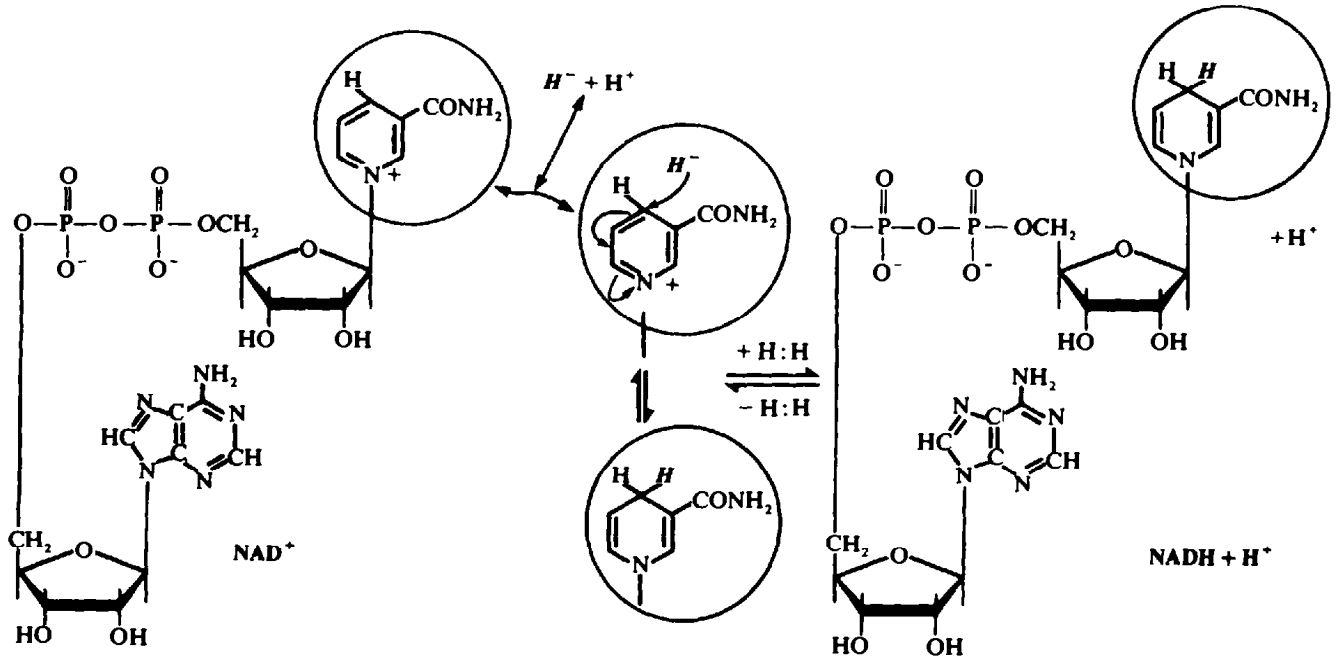


Fig. 11-1 The interconversion of NAD^+ and NADH.

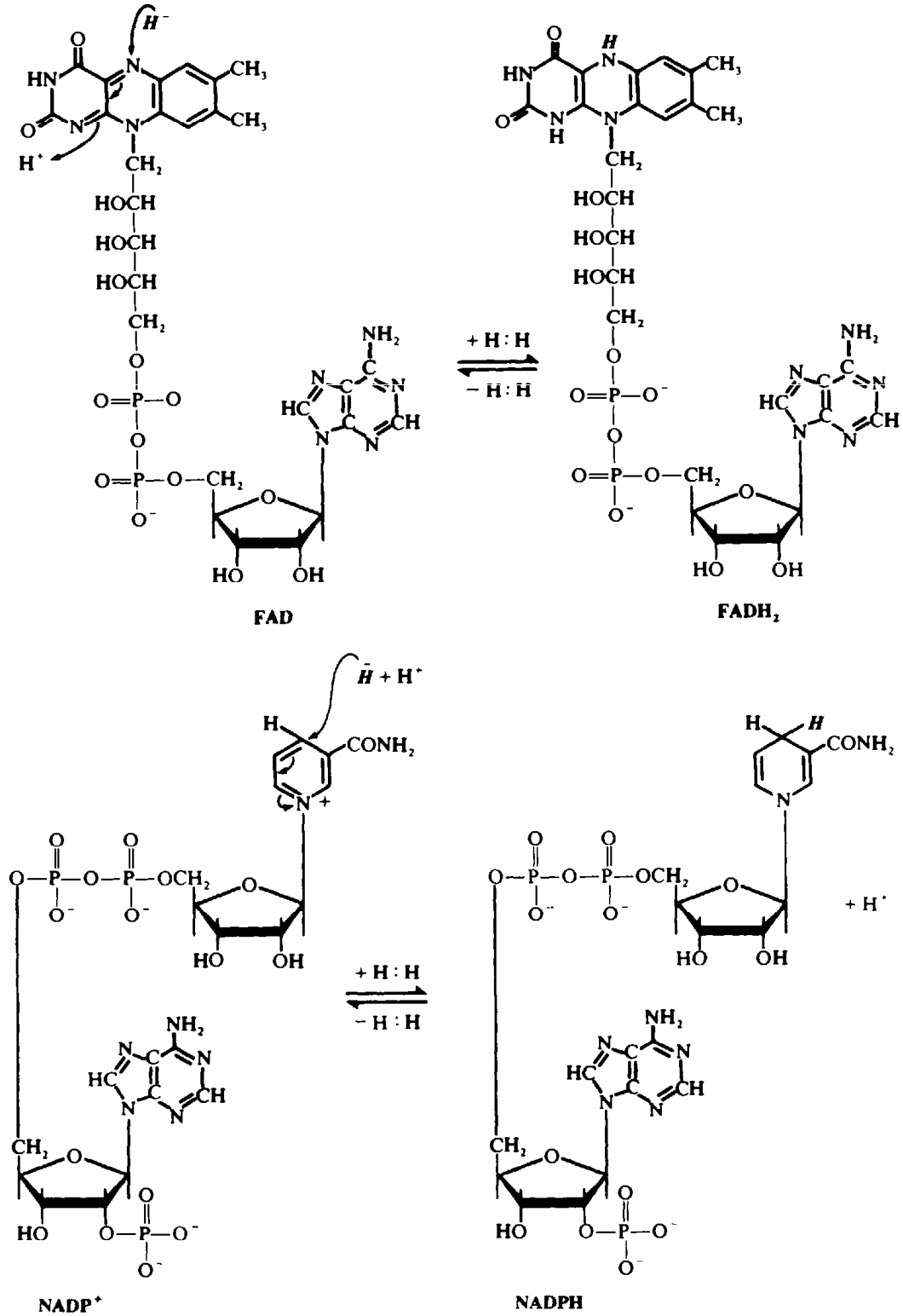


Fig. 11-2 The redox reactions of FAD/$FADH_2$ and $NADP^+$/NADPH.

EXAMPLE 11.1

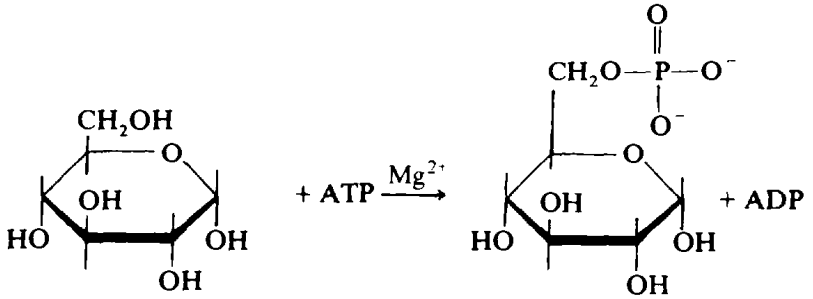
In the balanced chemical equation for glycolysis, two molecules of NAD^+ are converted to two molecules of NADH and two protons. The structure of NAD^+ (*n*icotinamide *a*denine *d*inucleotide) is given in Fig. 11.1. This is a reduction reaction, and the NAD^+, an enzyme cofactor, has accepted the equivalent of H^- (a *hydride*

ion). When a substrate is oxidized, it loses a pair of electrons that are contained within the structure H:H, and the function of NAD^+ is to accept these electrons. The relevant portion of the molecule involved in accepting the electrons as H^- is the aromatic pyridine ring. When the reduction has been completed, the reaction has produced NADH and H^+; just as NAD^+ can be *reduced* to NADH and H^+ by accepting 2H's, NADH and H^+ can be *oxidized* to NAD^+ by donating 2H's.

There are two other cofactors that can participate in *redox* processes; these are *f*lavin *a*denine *d*inucleotide (FAD) and *n*icotinamide *a*denine *d*inucleotide *p*hosphate ($NADP^+$), both of which are shown in Fig. 11-2. FAD accepts 2H's and is thereby reduced to $FADH_2$, whereas $NADP^+$ accepts H^- and is reduced to NADPH and H^+. Both of these reduced cofactors can be oxidized, thereby donating their H's (or *reducing equivalents*), similar to the oxidation of NADH. The enzymes that catalyze those reactions involving an oxidation or a reduction are usually very selective toward a particular cofactor (NAD or NADP) in a particular oxidation state.

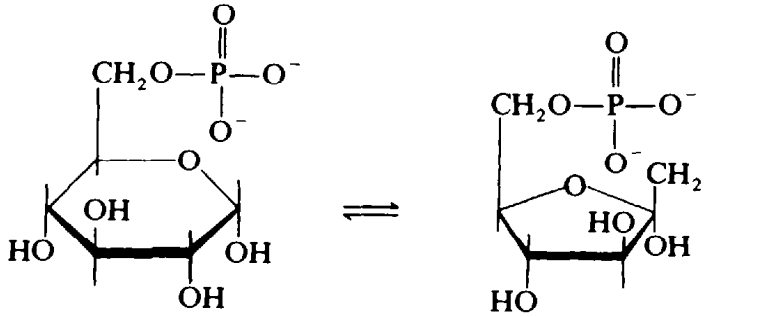
The steps in glycolysis are as follows:

Step 1 is the phosphorylation of α-D-glucose to α-D-glucose 6-phosphate. The enzyme catalyzing the reaction is *hexokinase*, and ATP and Mg^{2+} are required.



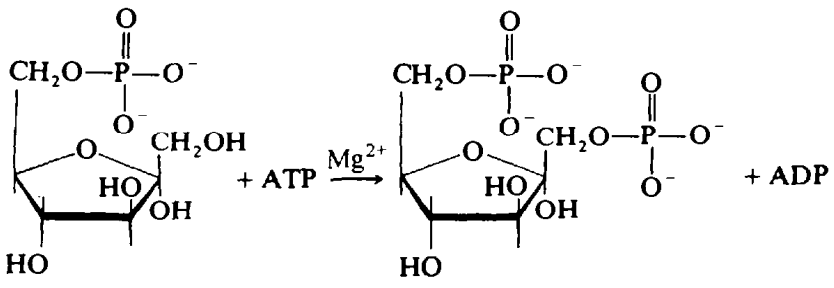
α-D-Glucose α-D-Glucose 6-phosphate

Step 2 is the isomerization of α-D-glucose 6-phosphate, by *glucose-6-phosphate isomerase*, to α-D-fructose 6-phosphate. This is a freely reversible reaction.



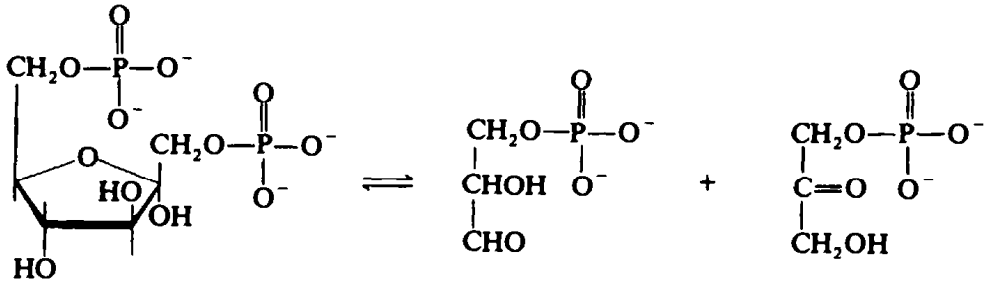
α-D-Glucose 6-phosphate α-D-Fructose 6-phosphate

Step 3 is the phosphorylation of α-D-fructose 6-phosphate to α-D-fructose 1,6-bisphosphate. The enzyme is *phosphofructokinase* and ATP and Mg^{2+} are required.



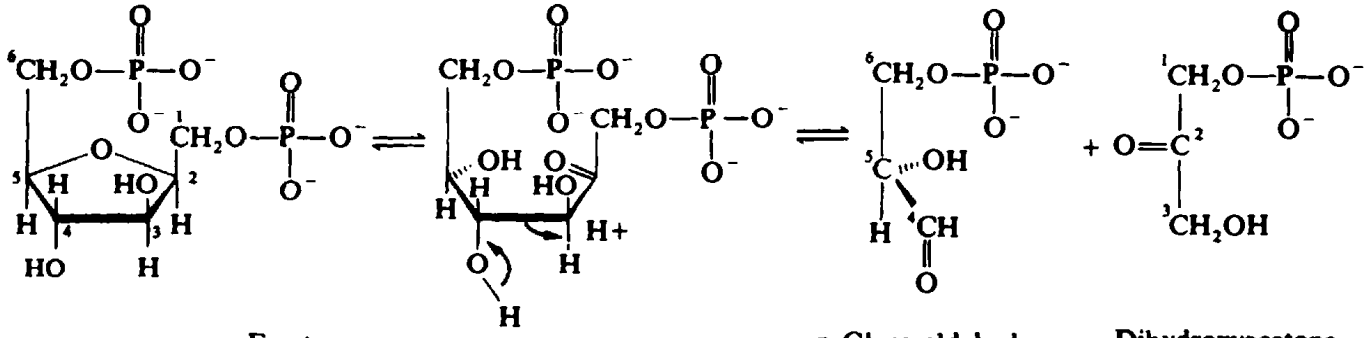
α-D-Fructose 6-phosphate α-D-Fructose 1,6-bisphosphate

Step 4 is the cleavage of α-D-fructose 1,6-bisphosphate. The reaction, which is reversible, is catalyzed by *fructose-1,6-bisphosphate aldolase*; this enzyme is usually referred to simply as *aldolase*.



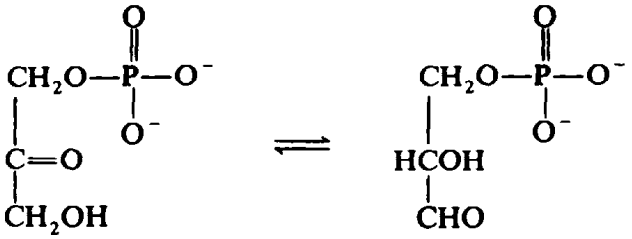
α-D-Fructose 1,6-bisphosphate | D-Glyceraldehyde 3-phosphate | Dihydroxyacetone phosphate

The mechanism of Step 4 involves a retro-aldol condensation reaction on the open-chain form of D-fructose 1,6-bisphosphate. This reaction, and the origin of the carbon atoms in the products, is shown below.



α-D-Fructose 1,6-bisphosphate | D-Glyceraldehyde 3-phosphate | Dihydroxyacetone phosphate

Step 5 is an isomerization. Dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate are interconverted by the enzyme *triosephosphate isomerase*.



Dihydroxyacetone phosphate | D-Glyceraldehyde 3-phosphate

Steps 4 and 5 result in the production of two molecules of D-glyceraldehyde 3-phosphate from one molecule of α-D-fructose 1,6-bisphosphate. The glycolytic pathway to this point (i.e., the conversion of glucose into two molecules of D-glyceraldehyde 3-phosphate) is called the *first stage* of glycolysis, and two molecules of ATP are required (at Steps 1 and 3) to provide the necessary energy. The remaining five steps compose the *second stage* of glycolysis and yield two molecules of ATP from ADP for *each* of the two three-carbon compounds produced above.

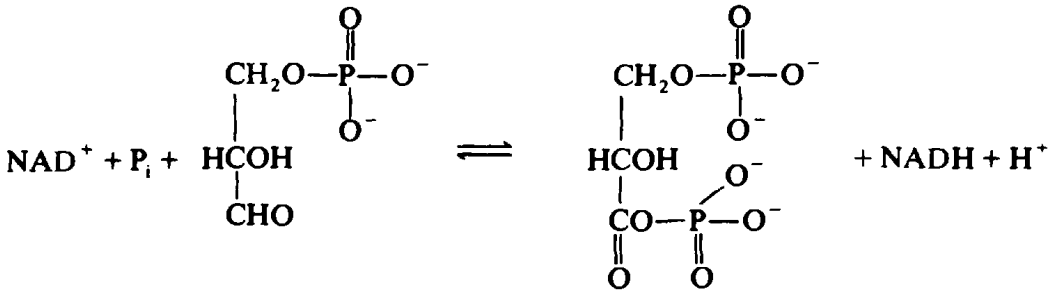
Question: How does the overall process of glycolysis

$$C_6H_{12}O_6 + 2ADP + 2NAD^+ + 2P_i \rightarrow 2C_3H_4O_3 + 2ATP + 2NADH + 2H^+ + 2H_2O$$

produce a net yield of two molecules of ATP for each molecule of glucose?

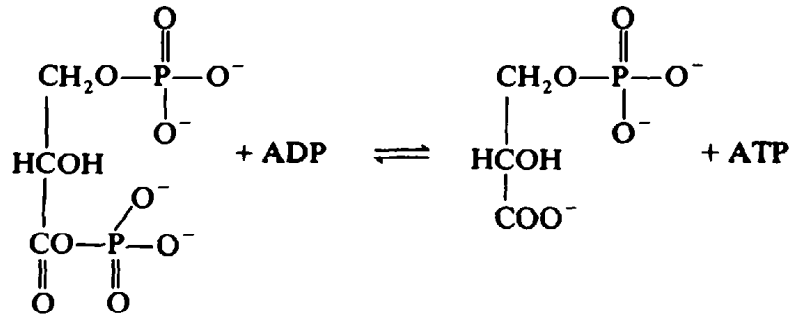
The first stage produces *two* molecules of D-glyceraldehyde 3-phosphate from one molecule of glucose. The second stage results in the formation of *two* molecules of ATP for *each* molecule of D-glyceraldehyde 3-phosphate used. Therefore, a net two molecules of ATP are synthesized from each molecule of glucose.

Step 6 involves the oxidation of D-glyceraldehyde 3-phosphate, accompanied by phosphorylation of the intermediate carboxylic acid, to produce D-1,3-bisphosphoglycerate. The enzyme is *glyceraldehyde-3-phosphate dehydrogenase.*



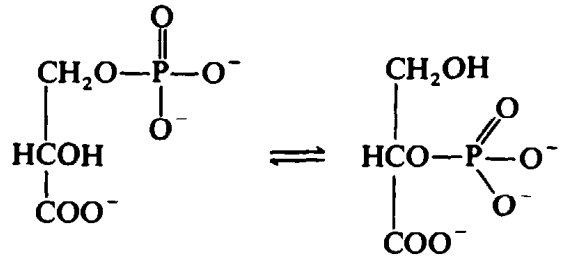
D-Glyceraldehyde 3-phosphate D-1,3-Bisphosphoglycerate

Step 7, the conversion of D-1,3-bisphosphoglycerate to D-3-phosphoglycerate, is an energy-yielding reaction, and ATP is produced. The enzyme is *phosphoglycerate kinase.*



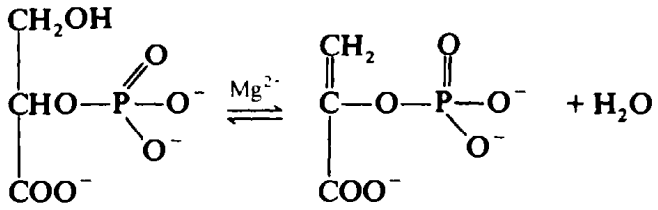
D-1,3-Bisphosphoglycerate D-3-Phosphoglycerate

Step 8 is an isomerization between D-3-phosphoglycerate to D-2-phosphoglycerate. This is catalyzed by the enzyme *phosphoglyceromutase.*



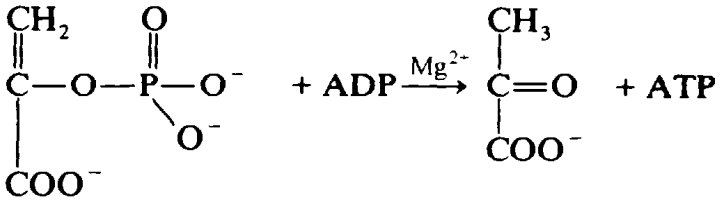
D-3-Phosphoglycerate D-2-Phosphoglycerate

Step 9 involves the dehydration of D-2-phosphoglycerate by *enolase* to produce phosphoenolpyruvate. The reaction also requires Mg^{2+}.



D-2-Phosphoglycerate **Phosphoenolpyruvate**

Step 10, the final step in glycolysis, is the irreversible conversion of phosphoenolpyruvate to pyruvate, catalyzed by *pyruvate kinase*. This is the second energy-yielding step in the glycolytic pathway, and produces ATP; Mg^{2+} is required here, too.



Phosphoenolpyruvate **Pyruvate**

Question: What do Steps 1, 3, and 10 of glycolysis have in common?

These are the only steps that are not freely reversible. They are the main control points in the pathway.

EXAMPLE 11.2

Several compounds can inhibit various enzymes in the glycolytic pathway. Enolase is inhibited by fluoride ions, and in vitro, glyceraldehyde-3-phosphate dehydrogenase is inactivated by iodoacetamide. In the presence of arsenate ions, glyceraldehyde 3-phosphate dehydrogenase can catalyze the conversion of glyceraldehyde 3-phosphate to phosphoglycerol arsenate, which hydrolyzes spontaneously to 3-phosphoglycerate and the arsenate ion. While this is not an *inhibition* of the enzyme, no ATP is produced at the phosphoglycerate kinase catalyzed reaction. This is one of the reactions affected in arsenic poisoning.

The Energetics of Glycolysis

The conversion of glucose to two molecules of pyruvate is an *exergonic* process (Chap. 10).

$$C_6H_{12}O_6 \rightarrow 2C_3H_4O_3 \qquad \Delta G^{\circ\prime} = -147\ \text{kJ mol}^{-1}$$

There is a sufficient decrease in the free energy of glucose to couple the process of glucose degradation with the substrate-level phosphorylation of two molecules of ADP, which is an *endergonic* reaction:

$$ADP + P_i \rightarrow ATP + H_2O \qquad \Delta G^{\circ\prime} = +30\ \text{kJ mol}^{-1}$$

Since the change in free energy required to produce two ATP molecules is +60 kJ, the coupled reaction of glucose → 2 pyruvate and 2ADP → 2ATP is:

$$C_6H_{12}O_6 + 2P_i + 2ADP \rightarrow 2C_3H_4O_3 + 2ATP + 2H_2O$$

$$\Delta G^{\circ\prime} = -147 + 60\ \text{kJ mol} = -87\ \text{kJ mol}^{-1}$$

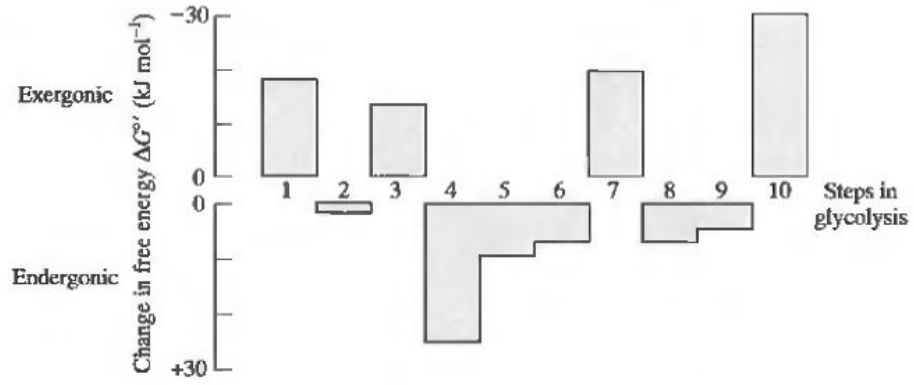


Fig. 11-3 Free energy of the glycolytic reactions.

This means that the conversion of glucose to pyruvate is still favored, even though it is coupled to the energy-requiring production of two molecules of ATP.

Four steps in glycolysis are exergonic and are, consequently, *favored* reactions (Fig. 11-3). All these steps involve either the hydrolysis of ATP (Steps 1 and 3) or the production of ATP (Steps 7 and 10). While the overall process favors the production of pyruvate, the individual changes in free energy are not particularly great. This suggests that it should be possible to reverse most of the steps without having to sacrifice large amounts of energy.

In the complete oxidation of glucose to CO_2:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

The standard free energy change is $-2{,}870\ \text{kJ mol}^{-1}$. When compared with the conversion of glucose to pyruvate ($\Delta G^{\circ\prime} = -87\ \text{kJ mol}^{-1}$), it is seen that only 3 percent of the energy available in the glucose molecule is released in producing pyruvate.

The Control Points of Glycolysis

The rate of glycolysis, as with all metabolic pathways, is under control; it is controlled at *three* stages.

The First Control Point

Step 1 is the first control point. In this step, glucose is converted to glucose 6-phosphate via hexokinase. This enzyme, which is present in all cells, is not specific for glucose but will catalyze the phosphorylation of many hexoses and hexose derivatives. However, the activity of the enzyme is regulated by the concentration of its principal product, *glucose 6-phosphate*; this product inhibits the activity of hexokinase in a process known as *product inhibition*. In mammalian cells, this regulatory function serves two purposes. First, it ensures that if a cell has sufficient glucose 6-phosphate to meet its energy demands, then subsequent phosphorylation of glucose will be reduced in that particular cell. Second, since the rate of the removal of glucose from the blood (by its conversion to glucose 6-phosphate within cells) is decreased, there will be an increase in the concentration of blood glucose if the glucose supply from elsewhere continues. The outcome of this is that the glucose will become more available to another phosphorylating enzyme, *glucokinase*. Glucokinase, which is *specific* for *D-glucose* and which is found only in the liver, also converts glucose to glucose 6-phosphate. Fig. 11-4 shows the relative activities of hexokinase (in all cells) and glucokinase (liver only).

Under normal conditions, glucose is available from the blood to all cells. The low K_m of hexokinase (0.1 mM) implies that even at *low* concentrations, glucose entering a cell is rapidly

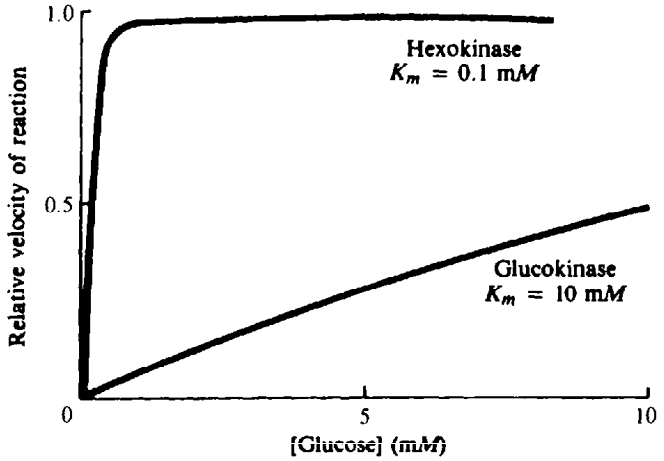


Fig. 11-4 The dependence of the rate of glucose phosphorylation on glucose concentration for hexokinase and glucokinase. The K_m for hexokinase is much lower than that for glucokinase.

converted to glucose 6-phosphate, which then enters the glycolytic pathway. As the energy requirements of the cell are met, the concentration of glucose 6-phosphate rises and thus reduces the activity of hexokinase. If the blood glucose concentration rises (e.g., after a carbohydrate-rich meal), then glucose flux through liver glucokinase increases. This occurs because hexokinase is fully saturated but glucokinase does not operate near its maximal rate until glucose levels rise beyond its K_m of 10 mM. Furthermore, glucokinase is not inhibited by glucose 6-phosphate. The interplay of these two enzymes ensures that if glucose is in excess of normal demands, it is converted into glucose 6-phosphate *specifically* within the liver.

The Second Control Point

Step 3 is the second control point of glycolysis and involves the conversion of fructose 6-phosphate into fructose 1,6-bisphosphate, catalyzed by phosphofructokinase.

Question: Which enzyme is the main point of control of glycolysis?

It is phosphofructokinase, which is an *allosteric* enzyme; thus, its activity is regulated by a number of effectors (Chap. 10), all of which are involved in energy transduction.

The activity of phosphofructokinase is enhanced by ADP or AMP and inhibited by ATP, NADH, citrate, or long-chain fatty acids. When a cell is in a *low*-energy state, the amounts of ADP and AMP are high relative to normal, whereas the amount of ATP is low. Under these conditions, the enzyme is fully activated and has a high affinity for its substrate, fructose 6-phosphate (Fig. 11-5). When the cell is in a high-energy state, the concentration of ATP is high, but the concentrations of AMP and ADP are low. In this circumstance, ATP binds to a regulatory site (Chap. 9) on the enzyme, causing its velocity curve to change from hyperbolic to sigmoidal. The enzyme now has a lower affinity for its substrate, and the rate of the reaction decreases.

Question: Citrate, NADH, and long-chain fatty acids inhibit the activity of phosphofructokinase. From where do these effectors arise, and what is the significance of their action?

Cytoplasmic NADH is produced in glycolysis at Step 6, so a high concentration of this *reduced* cofactor implies a high-energy state of the cell; thus, an increase in the rate of glucose degradation

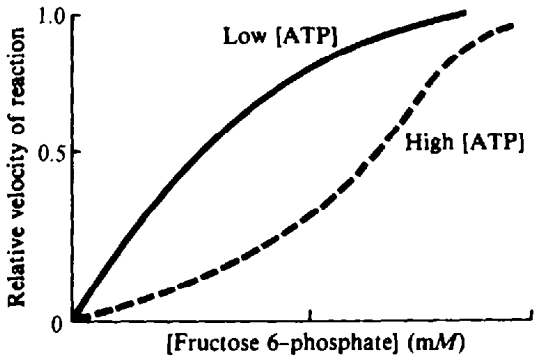


Fig. 11-5 The dependence of the rate of fructose 6-phosphate phosphorylation by phosphofructokinase on the presence of low and high concentrations of ATP.

does not occur. High amounts of long-chain fatty acids are produced by the degradation of triglycerides (Chap. 13) and, in the case of citrate, by the degradation of certain amino acids (Chap. 15). Since these substrates are available for oxidation in the citric acid cycle (Chap. 12), their effect in inhibiting phosphofructokinase is to conserve glucose.

The Third Control Point

Step 10 is the third control point of glycolysis. It involves the conversion of phosphoenolpyruvate to pyruvate, catalyzed by pyruvate kinase. This enzyme is *activated* by fructose 1,6-bisphosphate and phosphoenolpyruvate and is *inhibited* by ATP, citrate, and long-chain fatty acids. This means that the activity of pyruvate kinase is regulated in a manner similar to that of phosphofructokinase; both enzymes are inhibited when the cell is in a high-energy state or when alternative fuels to glucose are available. Furthermore, fructose 1,6-bisphosphate (the product of the reaction catalyzed by phosphofructokinase) activates pyruvate kinase, as does phosphoenolpyruvate (the substrate of pyruvate kinase). These are examples of *positive feedforward control* akin to some electronic circuitry. So, when phosphofructokinase is activated by low levels of ATP, it produces one activator for pyruvate kinase (fructose 1,6-bisphosphate) that is ultimately converted into a second activator of pyruvate kinase (phosphoenolpyruvate). This *cooperation* between the two enzymes to accelerate glycolysis also extends to their combined ability to retard the process. When the concentration of ATP is high, both enzymes are inhibited. Phosphofructokinase has reduced activity toward fructose 6-phosphate, so the concentration of this component rises, and because it interconverts to glucose 6-phosphate (via glucose 6-phosphate isomerase), the concentration of this substrate increases, thereby inhibiting hexokinase (Fig. 11-6).

Not only is glycolysis in all cells controlled by these three enzymes, but if the cell is in a high-energy state, or if glucose is plentiful, or both, the excess glucose is not degraded by other tissues but is selectively captured via glucokinase in the liver, where it is stored (Sec. 11.5).

11.2 THE FATE OF PYRUVATE

The production of two molecules of pyruvate from one molecule of glucose occurs in virtually all cells. This process has three important characteristics: (1) no oxygen is required; (2) two molecules of ADP are phosphorylated by substrate-level phosphorylation; and (3) two molecules of NAD^+ are reduced. The subsequent fate of pyruvate in a particular cell depends on conditions related to these

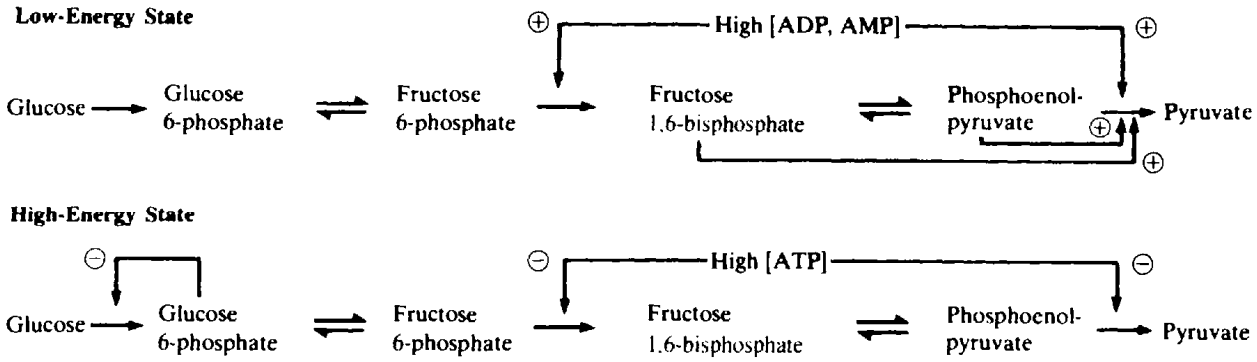


Fig. 11-6 The controls of glycolysis. ⊕ = activation; ⊖ = inhibition.

three criteria. The first is oxygen availability to the cell; the second is the energy status of the cell; and the third concerns the mechanisms available to the cell to oxidize the NADH to NAD^+.

Question: Why must the NADH produced in glycolysis be oxidized to NAD^+ and thus be recycled?

NAD^+ is required by glyceraldeyde 3-phosphate dehydrogenase (Step 6); therefore, NAD^+ is essential for this step and for glycolysis to proceed. Cells do not contain high concentrations of either NAD^+ or NADH; there must be mechanisms available for their interconversion so that each is available for the reactions in which it takes part.

A further criterion governing the fate of pyruvate is the type of cell in which it is formed, since some cells (e.g., red blood cells) lack the metabolic capability to carry out the complete oxidation of pyruvate to CO_2.

Complete Oxidation of Pyruvate to CO_2

On complete oxidation, a mole of glucose liberates 2,870 kJ of free energy. In its conversion to two moles of pyruvate, only 3 percent of this available energy is released and only two moles of ATP are produced. The majority of the remaining energy is available to the cell only if it has the capacity to oxidize the pyruvate completely to CO_2.

The chemical equation for the complete oxidation of pyruvate is

$$C_3H_4O_3 + 2\tfrac{1}{2}O_2 \rightarrow 3CO_2 + 2H_2O$$

This equation actually represents two oxidative processes, the first being the oxidation of pyruvate to CO_2 in the citric acid cycle (Chap. 12):

$$C_3H_4O_3 + 3H_2O \rightarrow 3CO_2 + 10H$$

This process, in which pyruvate yields the equivalent of 10 H's, also results in the direct phosphorylation of 1 molecule of ADP. The second process is

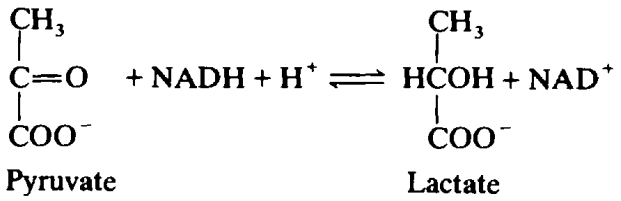
$$10H + 2\tfrac{1}{2}O_2 \rightarrow 5H_2O$$

It describes the oxidation of the 10 H's in the electron-transport chain (Chap. 14), with the concomitant phosphorylation of 14 molecules of ADP by oxidative phosphorylation. Consequently, by being completely oxidized to CO_2, 1 molecule of pyruvate leads to the phosphorylation of 15

molecules of ADP. When glucose is degraded to 2 molecules of pyruvate, only 2 molecules of ATP are formed, but the 2 pyruvate molecules have the potential to generate a further 30 molecules of ATP. The ability to produce this ATP depends on two factors: (1) the cell must have the capacity to perform both the citric acid cycle and electron transport, and (2) it must have a supply of oxygen. If either of these two criteria is lacking, then the 30 molecules of ATP cannot be produced. One example of a mammalian cell that lacks the ability to perform the citric acid cycle and electron transport is the red blood cell; it lacks this ability because it does not possess mitochondria. Skeletal muscle, when it is very active, has a limited capacity to oxidize pyruvate completely, because it has far fewer mitochondria than most other cells, and the amount of pyruvate produced during muscular activity cannot be fully oxidized.

Conversion of Pyruvate to Lactate

If a cell lacks the ability to oxidize pyruvate, it is restricted to the glycolytic process for its production of ATP. If sufficient glucose is available to the cell, pyruvate is disposed of, so long as ADP, NAD^+, and P_i are present. All cells have adequate amounts of ADP and P_i, since these are hydrolysis products of ATP, but the amounts of NAD^+ are more limited. Step 6 is the only oxidative reaction in glycolysis; glyceraldehyde 3-phosphate is oxidized to 1,3-bisphosphoglycerate, while NAD^+ is reduced to NADH. For continued glycolysis, this NADH must be reoxidized to NAD^+. This occurs in red blood cells and active muscle cells by the reduction of pyruvate to give lactate, which diffuses out of the cell via a specific membrane transport protein. The reaction is catalyzed by *lactate dehydrogenase*:



Pyruvate **Lactate**

All mammalian cells possess lactate dehydrogenase, but the activity of the enzyme varies from tissue to tissue. This variation is due to there being five forms of the enzyme, called *isozymes* or *isoenzymes*, each possessing a different "apparent" K_m for pyruvate. Each isozyme of lactate dehydrogenase consists of four subunits of either type M or H. So the five isozymes of lactate dehydrogenase are M_4, M_3H, M_2H_2, MH_3, and H_4. Isozyme M_4 has a relatively high K_m (low affinity) for pyruvate but has a turnover number approximately twice that of H_4. This means that at high concentrations of pyruvate, M_4 will convert pyruvate to lactate at a faster rate than will an equivalent amount of isozyme H_4. Even though H_4 has a low K_m (high affinity) for pyruvate, it has a *lower* turnover number so that at high concentrations of pyruvate it will convert pyruvate to lactate at a slower rate than will an equivalent amount of M_4. Furthermore, isozyme H_4 is inhibited by high concentrations of pyruvate. The other isozymes have degrees of inhibition intermediate to these two extremes. All cells possess varying amounts of the five isozymes; skeletal muscle, for example, has a predominance of the M_4 isozyme, while heart muscle has a predominance of H_4. The fact that heart muscle cells contain a predominance of the isozyme with the lowest turnover number for pyruvate and highest inhibition by pyruvate is consistent with the tendency of these cells to oxidize pyruvate to CO_2 and not, as is the case of active skeletal muscle, to convert it to lactate.

Although liver cells can oxidize pyruvate to CO_2, they contain a predominance of the M_4 isozyme, with its low affinity for pyruvate. Lactate enters the liver from the blood plasma and is rapidly converted to pyruvate.

This metabolic scheme, which is called *lactate fermentation*, is shown in Fig. 11-7. The coreactant cycle between the two dehydrogenase enzymes, glyceraldehyde-3-phosphate dehydrogenase (Step 6) and lactate dehydrogenase, ensures that there is regeneration of NAD^+ in this particular oxidation state so that glycolysis, lactate fermentation, and the production of ATP can continue.

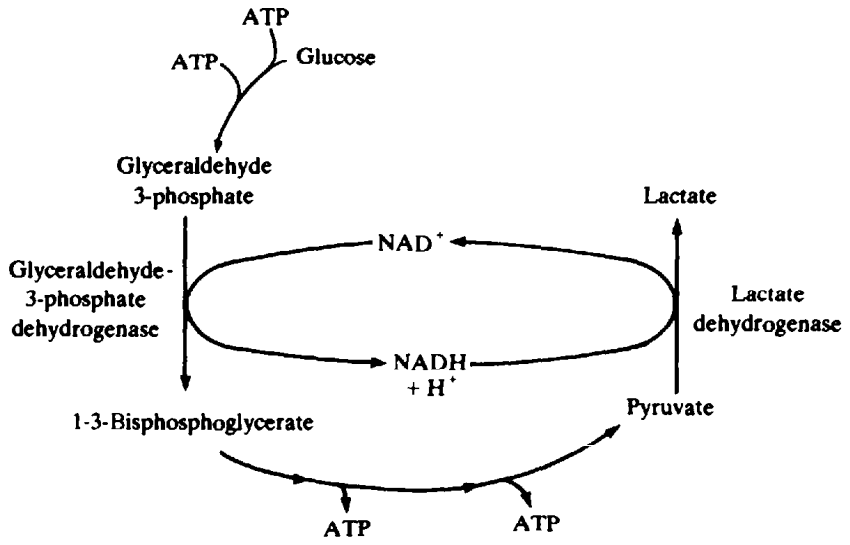
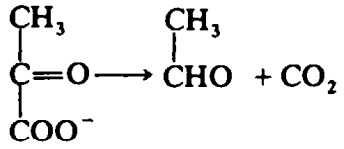


Fig. 11-7 The cooperation between glyceraldehyde 3-phosphate dehydrogenase and lactate dehydrogenase.

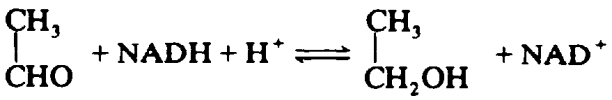
Conversion of Pyruvate to Ethanol

There is another fate of pyruvate that, while not occurring in mammalian tissues, is nevertheless very important. Some organisms can live under aerobic or anaerobic conditions. These are called *facultative anaerobes*, and they can alter their metabolism to adapt to the presence or absence of oxygen. The most important facultative anaerobes are the yeasts. They convert glucose to pyruvate and then, if oxygen is present, oxidize the pyruvate to CO_2. If there is no oxygen available, then a pathway for the regeneration of NAD^+ comes into operation. Yeasts do not have lactate dehydrogenase but do possess *pyruvate decarboxylase*, which is not present in mammalian cells. This enzyme catalyzes the conversion of pyruvate to *acetaldehyde*:



Pyruvate **Acetaldehyde**

Acetaldehyde is then reduced to ethanol (ethyl alcohol) via the Zn-containing enzyme *alcohol dehydrogenase*:



Acetaldehyde **Ethanol**

This, the final step in alcohol fermentation, is analogous to lactate fermentation. Both reactions regenerate NAD^+ and produce low-molecular-weight, water-soluble, metabolic end products that diffuse out of the cells in which they were produced. In the case of alcoholic fermentation, the second reaction is reversible, so that if oxygen becomes available to previously anaerobic yeast cells, the ethanol is oxidized to acetaldehyde. Unlike lactate fermentation, in which the lactate is oxidized to pyruvate, alcoholic fermentation cannot form pyruvate from acetaldehyde. Instead, the acetaldehyde

is converted to acetic acid, which is then completely oxidized to CO_2 in the citric acid cycle. The oxidation of acetic acid yields eight H's, which generate ATP in the process of *oxidative phosphorylation* (Chap. 14):

$$CH_3COOH + 2H_2O \rightarrow 2CO_2 + 8H$$

$$8H + 2O_2 \rightarrow 4H_2O$$

The overall reaction is:

$$CH_3COOH + 2O_2 \rightarrow 2CO_2 + 2H_2O$$

11.3 GLUCONEOGENESIS

Question: It has been stated above that tissues can break down glucose, but can they synthesize it too?

Yes, but only certain tissues have this ability.

In mammalian cells, glucose is the most abundant carbohydrate energy source. It is metabolized in all cells as a glycolytic fuel and is stored in liver and muscle as the polymer glycogen. But certain cells have the enzymes to catalyze the synthesis of glucose under certain conditions. The requirements are (1) the availability of *specific* carbon skeletons (carbon backbone structures of various types), (2) energy, in the form of ATP, necessary to accomplish the sequence of reactions, and (3) the enzymes to catalyze reactions of the sequence.

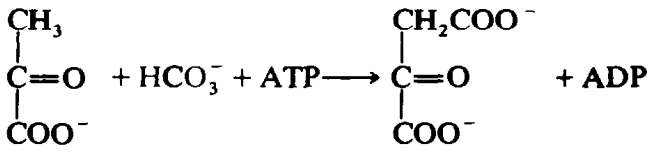
The carbon skeletons used for the synthesis of glucose are *not* of *carbohydrate* origin but are derived from particular amino acids. One exception to this is the carbon skeleton of lactate, itself a product of carbohydrate metabolism, which can be incorporated into a new glucose molecule. This process, the synthesis of new glucose from essentially noncarbohydrate precursors, is called *gluconeogenesis*.

Question: Can the glycolytic pathway operate in the reverse direction; that is, can pyruvate be converted into glucose?

The pathway cannot operate *directly* in reverse because of the three irreversible steps, but pyruvate can be converted into glucose because of additional reactions.

The three steps in glycolysis that are irreversible are those catalyzed by hexokinase (glucokinase in the liver), phosphofructokinase, and pyruvate kinase. However, those tissues that carry out gluconeogenesis (e.g., liver and kidney) possess enzymes that allow these three steps to be reversed. When this occurs, the net flux of carbon atoms through the other reactions of glycolysis is reversed. A summary of the process is given in Fig. 11-8. Three nonglycolytic steps occur when pyruvate is converted to glucose via the *net reversal* of glycolysis.

Step A, the conversion of pyruvate to phosphoenolpyruvate, is accomplished by a circuitous process commencing with pyruvate entering the mitochondrion, which for gluconeogenesis to occur must be in a high-energy state. Under these conditions, the mitochondrial enzyme *pyruvate carboxylase* catalyzes the conversion of pyruvate to *oxaloacetate*:



Pyruvate **Oxaloacetate**

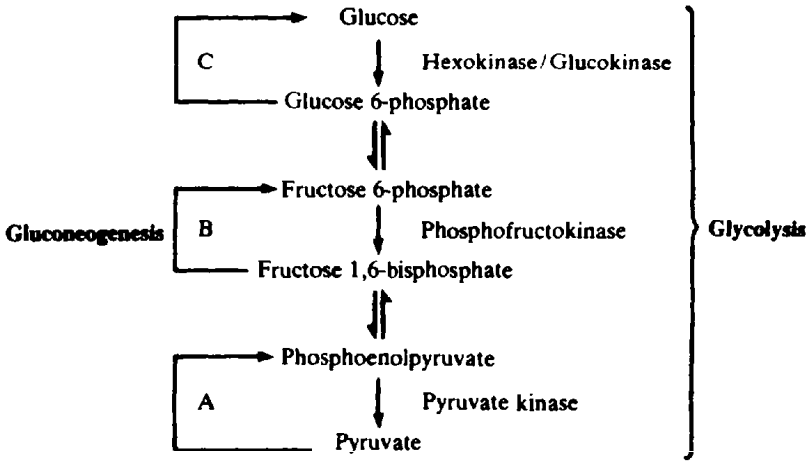
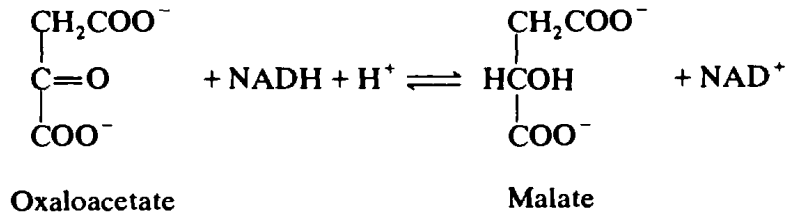
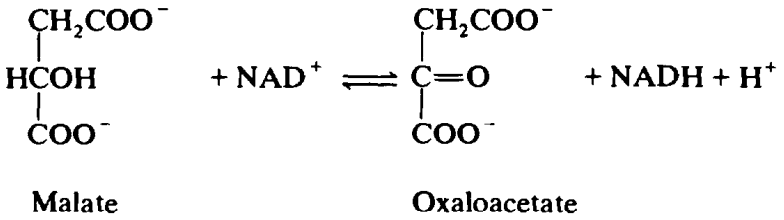


Fig. 11-8 Gluconeogenesis and glycolysis. A, B, and C denote steps in gluconeogenesis that bypass irreversible glycolytic reactions (in the reverse direction to that of net glycolytic flux).

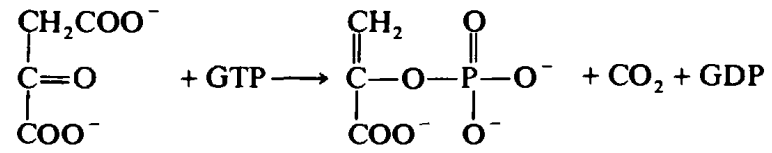
Pyruvate carboxylase is an allosteric enzyme being activated by its effector, acetyl-CoA. The enzyme contains tightly bound Mn^{2+} ions, and a covalently attached prosthetic group, biotin. When the mitochondria are in a high-energy state, the concentrations of acetyl-CoA and ATP are relatively high, so the modulator of the enzyme and a source of energy are both available. The oxaloacetate is then converted within mitochondria to *malate*:



The next reaction occurs in the cytoplasm. Malate is transported to the cytoplasm by a *dicarboxylate carrier* which is specific for malate, succinate, and fumarate and which requires the entry of P_i or one of these dicarboxylate anions. Cytoplasmic malate is then converted to oxaloacetate by *cytoplasmic malate dehydrogenase*:



The preceding two reactions are necessary to achieve the overall result of transporting oxaloacetate to the cytoplasm from the mitochondria, as there is no direct mechanism for this. Cytoplasmic oxaloacetate is then converted irreversibly to phosphoenolpyruvate by way of *phosphoenolpyruvate carboxykinase*, a cytoplasmic enzyme that operates only when the ATP concentration is high:



Oxaloacetate Phosphoenolpyruvate

These four reactions convert pyruvate to phosphoenolpyruvate (Step A, Fig. 11-8) and bypass the irreversible step in glycolysis that is catalyzed by pyruvate kinase (Step 10).

Reversal of glycolytic Steps 9 through to 4 is achieved at the expense of supplying energy, to Step 7, in the form of ATP. In other words, net gluconeogenesis occurs only under conditions of relatively high cellular free energy. The high-energy state also implies that the NADH concentration is high; this enables the reversal of Step 6 in glycolysis, which involves the conversion of 1,3-bisphosphoglycerate to glyceraldehyde 3-phosphate via glyceraldehyde-3-phosphate dehydrogenase (Fig. 11-9).

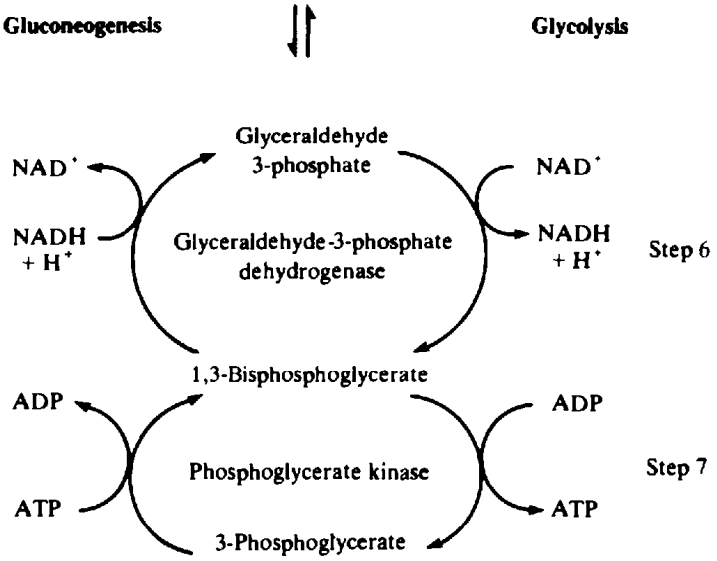
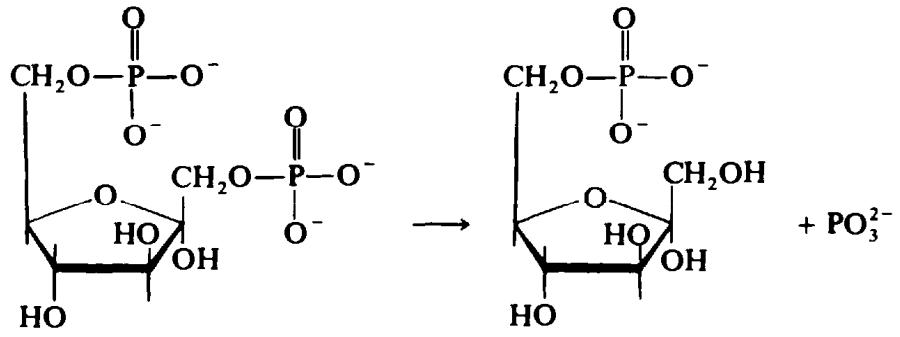


Fig. 11-9 The reversal of Steps 6 and 7 in glycolysis is favored by a high-energy state.

As a result of the reversal of Step 4 in glycolysis, the equivalent of two molecules of pyruvate is condensed to give one molecule of fructose 1,6-bisphosphate. This compound is the product of the irreversible Step 3 in glycolysis. Gluconeogenic cells have the enzyme *fructose-1,6-bisphosphatase*, which catalyzes the reverse reaction (Step 3, p. 313).



Fructose 1,6-bisphosphate Fructose 6-phosphate

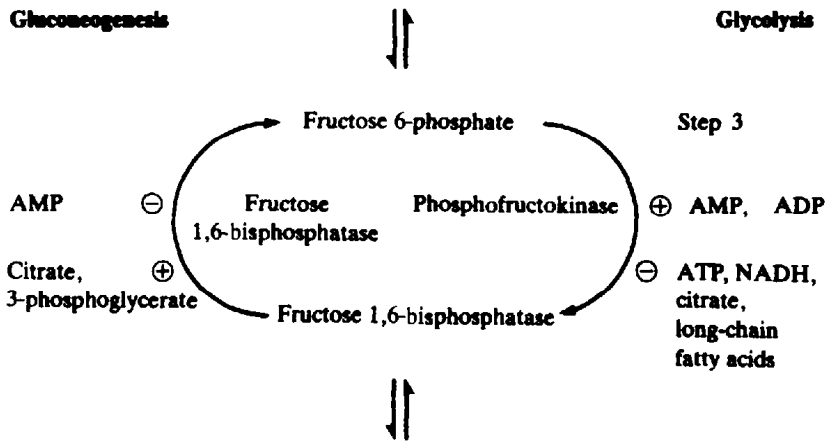
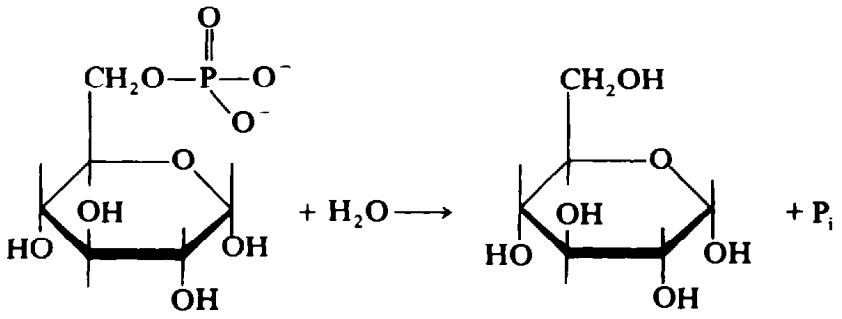


Fig. 11.10 The reversal of Step 3 in glycolysis is favored by a high-energy state. ⊕ = activation; ⊖ = inhibition.

Fructose 1,6-bisphosphatase is inhibited by AMP but activated by citrate and 3-phosphoglycerate. Thus, in a high-bond-energy state, an increase of citrate and a decrease of AMP combine to activate fructose-1,6-bisphosphatase and to inhibit phosphofructokinase (Fig. 11-10). This promotes the hydrolysis of fructose 1,6-bisphosphate to fructose 6-phosphate.

Fructose 6-phosphate is converted to glucose 6-phosphate via glucose-6-phosphate isomerase. Glucokinase (liver only) and hexokinase, the enzymes that produce glucose 6-phosphate from glucose, are not able to catalyze the reverse reaction, but the liver has a specific enzyme, *glucose-6-phosphatase*, that catalyzes the hydrolysis of glucose 6-phosphate to glucose. The glucose that is produced can then enter the blood.



Glucose 6-phosphate Glucose

An alternative metabolic route for glucose 6-phosphate is its storage as glycogen in the liver and muscles. (The production of glycogen is discussed in Sec. 11.5.)

In summary, glucose can be synthesized in the liver and kidney from lactate and noncarbohydrate precursors (carbons from certain amino acids). The synthesis is accomplished essentially by a reversal of the glycolytic pathway, because the cells of these tissues possess the enzymes that are necessary to overcome the three irreversible glycolytic steps.

11.4 THE CORI CYCLE

The localization of particular enzymes in only certain cells means that some organs depend on others for the complete metabolism of certain substrates. So far as carbohydrates are concerned, the

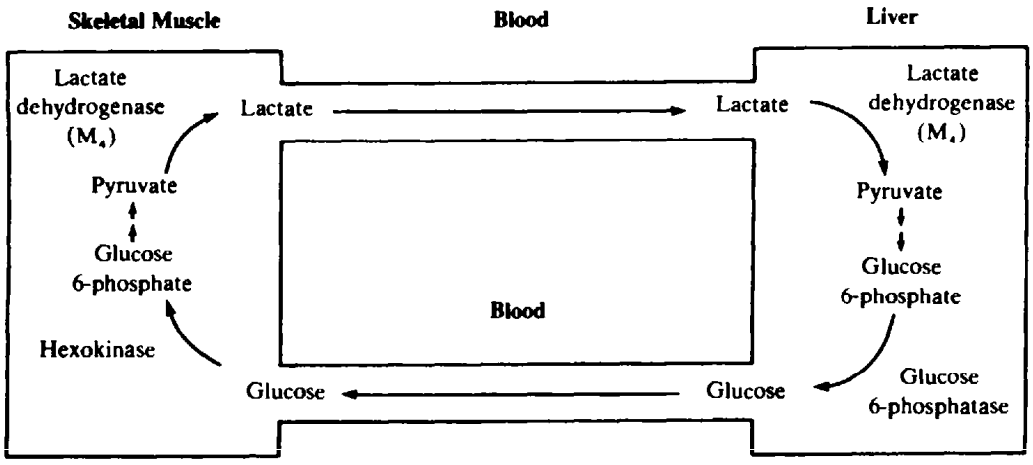


Fig. 11-11 The Cori cycle.

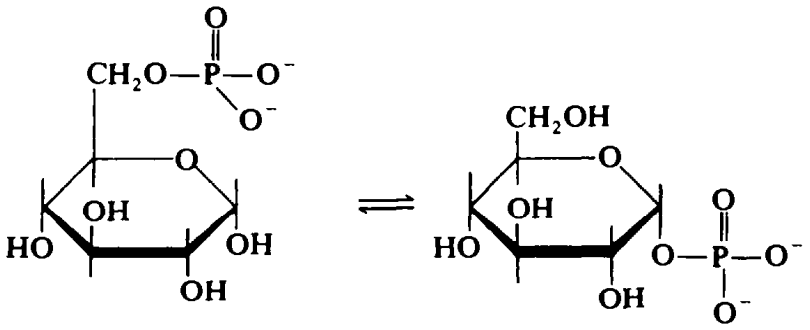
liver and skeletal muscle exhibit a special metabolic cooperation. Skeletal muscle obtains ATP during exercise almost solely from glycolysis. As a result, the end product, lactate, enters the blood. The lactate is removed from the blood by the liver mainly via the M_4 isozyme of lactate dehydrogenase, which catalyzes the rapid conversion of lactate into pyruvate. As the liver is usually in a high-energy state, the majority of this pyruvate is converted by the gluconeogenic pathway to glucose 6-phosphate. This can be hydrolyzed to glucose via glucose 6-phosphatase and can then enter the blood, where it is transported to skeletal muscles. In skeletal muscle the glucose is converted to glucose 6-phosphate via hexokinase and enters glycolysis. This process (Fig. 11-11) is called the *Cori cycle* after its discoverers, Carl and Gerty Cori, who were Nobel prize winners in 1937.

11.5 GLYCOGEN METABOLISM

The Synthesis of Glycogen

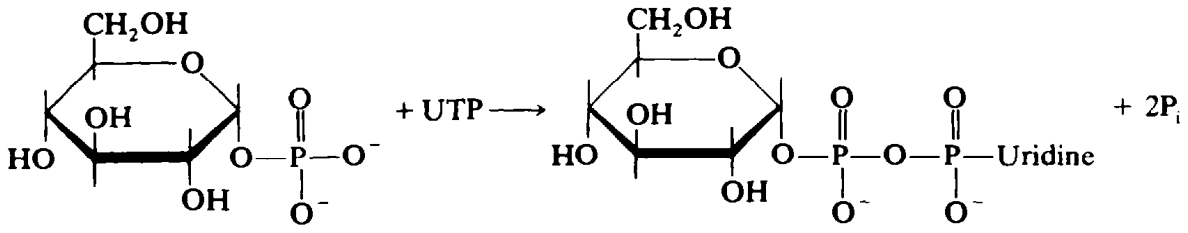
Glycogen is synthesized from glucose 6-phosphate in the liver and muscle and is stored within these tissues as *glycogen granules*. The glycogen, which is a polymer of glucose, is an energy store that can be rapidly broken down ultimately to glucose 6-phosphate, which then enters the glycolytic pathway. While the mechanism of synthesis of glycogen in all tissues is identical, the origin of the glucose 6-phosphate needs to be considered. In the liver, glucose 6-phosphate can arise either from blood glucose or through gluconeogenesis. In skeletal muscle, the glucose 6-phosphate originates *solely* from blood glucose.

The first step in glycogen synthesis is the formation of glucose 1-phosphate, catalyzed by *phosphoglucomutase:*



Glucose 6-phosphate Glucose 1-phosphate

The glucose 1-phosphate is then *activated* to enable its incorporation into glycogen. This activation involves the expenditure of energy derived from the hydrolysis of a molecule of *uridine triphosphate* (UTP) by *UDP-glucose pyrophosphorylase*.



Glucose 1-phosphate Uridine diphosphoglucose (UDP-glucose)

Question: What is the origin of the two phosphate groups in UDP-glucose?

One is from glucose 1-phosphate and the other is from *uridine monophosphate* (UMP). The pyrophosphate that is liberated from the terminal phosphates of UTP is hydrolyzed to inorganic phosphate by the enzyme *pyrophosphatase*. This hydrolysis, which is *irreversible*, drives the reaction in the direction of UDP-glucose synthesis.

The modified glucose molecule is a substrate for the enzyme *glycogen synthase*:

$$(\text{Glucose})_n + \text{UDP-glucose} \rightarrow (\text{Glucose})_{n+1} + \text{UDP}$$

The addition of a glucose unit to a polymeric chain of glucose is specific; i.e., the ester linkage forms only between the hydroxyl group on the C-4 of a terminal glucose unit of glycogen and the oxygen of the C-1 of the α isomer of the incoming UDP-glucose molecule (Fig. 11-12). The bond formed is, therefore, an $\alpha(1\rightarrow 4)$ linkage (Chap. 2). In addition, glycogen has branched chains joined in $\alpha(1\rightarrow 6)$ linkages. The enzyme *amylo*-(1,4→ 1,6)-*transglycosylase* catalyzes the transfer of fragments of six or seven glucose residues in glycogen to a C-6 hydroxyl group on another glucose residue within the glycogen polymer (Fig. 11-13).

The branching within glycogen molecules makes the polymers more *compact* and more *soluble* and produces more *terminal glucose residues*. The increase in terminal residues is important when glycogen is to be degraded, because this occurs through a stepwise cleavage of terminal glucose residues from the polymer.

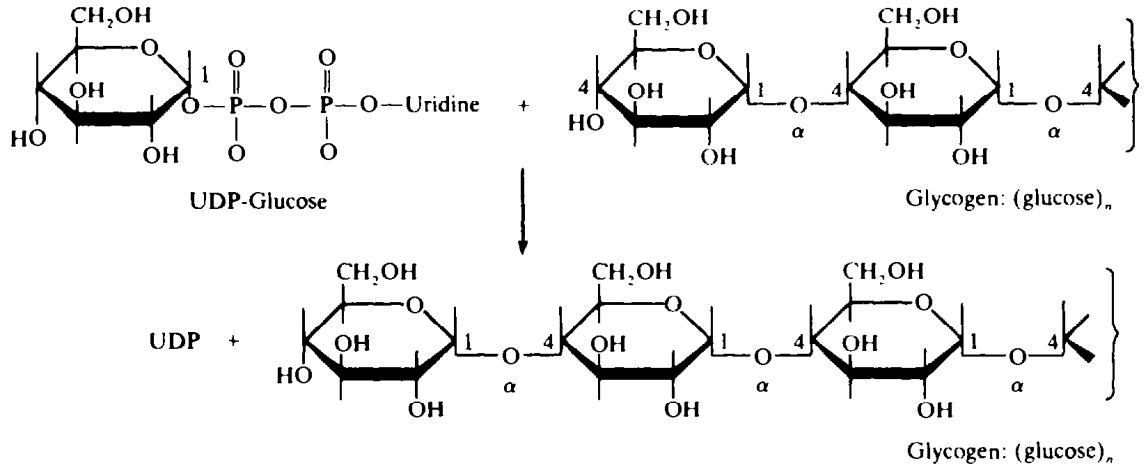


Fig. 11-12 The addition of a glucose unit to glycogen.

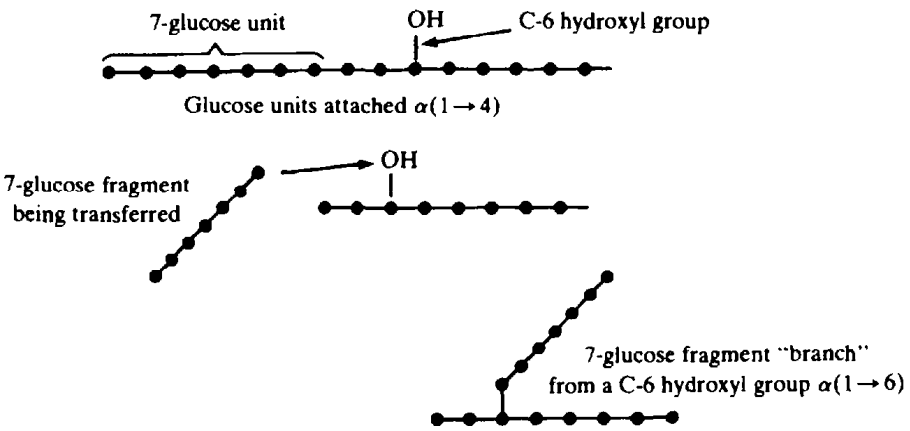
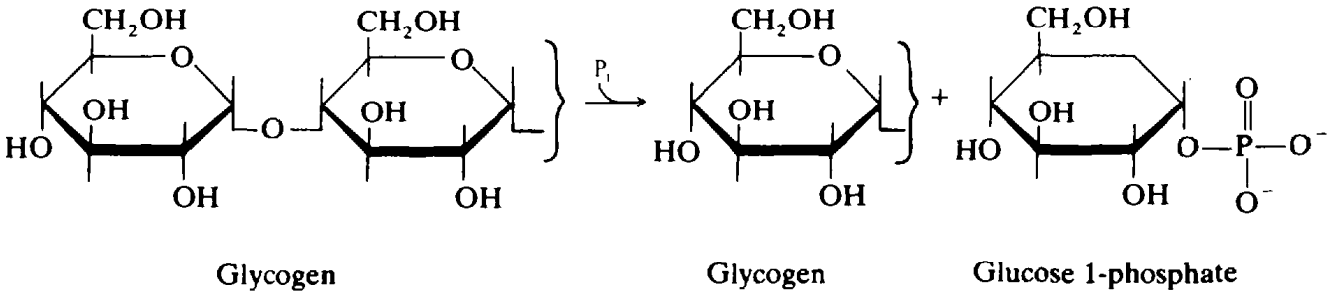


Fig. 11-13 The formation of branched chains of glucose residues within glycogen. ●— = glucose residue.

The Degradation of Glycogen (Glycogenolysis)

Glycogen is degraded to glucose 6-phosphate by a pathway that differs from its synthesis. The first step is via the enzyme *glycogen phosphorylase*, which, with inorganic phosphate, catalyzes the cleavage of a terminal $\alpha(1\rightarrow4)$ bond, provided an $\alpha(1\rightarrow6)$ linkage is not attached, to produce glycogen with one residue less and a molecule of glucose 1-phosphate.



Glycogen phosphorylase cannot cleave $\alpha(1\rightarrow6)$ linkages. This is carried out by another enzyme, called an $\alpha(1\rightarrow6)$-glucosidase, which hydrolyzes these bonds and thus makes more $\alpha(1\rightarrow4)$ linkages accessible to the actions of glycogen phosphorylase. Production of glucose 1-phosphate is followed by its conversion via the enzyme *phosphoglucomutase* to glucose 6-phosphate. The fate of the glucose 6-phosphate depends on whether it is formed in a skeletal muscle cell or in a liver cell. In skeletal muscle, the compound continues along the glycolytic pathway; while this also occurs in the liver, glucose 6-phosphatase can also convert glucose 6-phosphate to glucose.

Control of Glycogen Synthesis and Degradation

There are a number of factors controlling the synthesis and degradation of glycogen. The two enzymes concerned in these processes, *glycogen synthase* and *glycogen phosphorylase*, are allosterically controlled, and their activities are modulated by glucose 6-phosphate; when the concentration of glucose 6-phosphate is high, glycogen synthase is activated, thus favoring glycogen synthesis over glycogenolysis. A high concentration of glucose 6-phosphate inhibits the activity of glycogen phosphorylase, so that the degradation of glycogen is inhibited (Fig. 11-14). Another control in glycogenolysis is mediated by the hormones *epinephrine* and *glucagon* (Sec. 11.9).

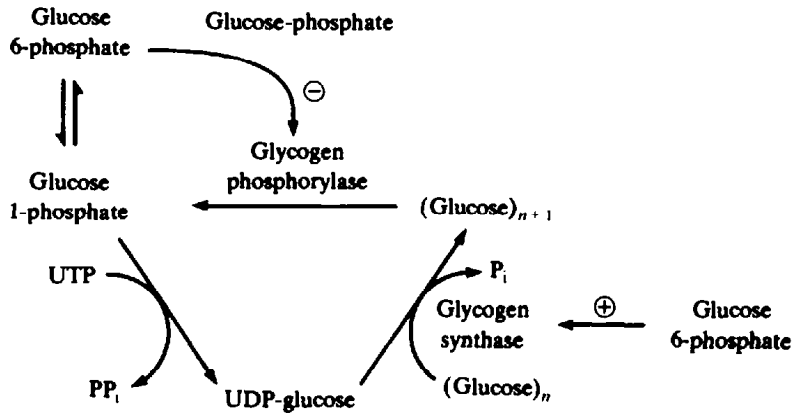


Fig. 11-14 The synthesis and degradation of glycogen. ⊕ = activation; ⊖ = inhibition.

11.6 THE ENTRY OF OTHER CARBOHYDRATES INTO GLYCOLYSIS

Glycolysis was defined as the process of converting one molecule of glucose into two molecules of pyruvate. However, many carbohydrates can contribute, to the cell, their carbon skeletons and the bond energy contained within their structures via the glycolytic pathway. In so doing, these other carbohydrates can ultimately be converted to pyruvate or to glucose and therefore be stored as glycogen.

EXAMPLE 11.3

The degradation of glycogen to glucose 1-phosphate is an example of the entry of a polysaccharide into glycolysis, since glucose 1-phosphate can be converted into glucose 6-phosphate. Another polysaccharide that can contribute its carbohydrate units to glycolysis is starch. *Starch* is the storage form of glucose in plants and has a structure similar to that of glycogen. It is a polymer of glucose units joined in $\alpha(1\rightarrow 4)$ and $\alpha(1\rightarrow 6)$ linkages but has fewer $\alpha(1\rightarrow 6)$ linkages than glycogen. Starch is really a mixture of two types of glucose polymers: an unbranched form called *amylose* and a branched form called *amylopectin*. During digestion, starch is hydrolyzed ultimately to maltose and glucose. Maltose is a *disaccharide* of two glucose units, joined by an $\alpha(1\rightarrow 4)$ link, and this is cleaved via the enzyme *maltase* to produce two glucose molecules. Consequently, the digestion of starch leads to the formation of glucose.

EXAMPLE 11.4

Other disaccharides commonly ingested by humans are *sucrose* (cane sugar) and *lactose* (milk sugar). Sucrose is cleaved into glucose and fructose by the action of *sucrase*, and lactose is cleaved into glucose and galactose by the action of *lactase*. All the monosaccharides can produce glycolytic intermediates. Glucose is converted via hexokinase into glucose 6-phosphate (Step 1 in glycolysis); fructose also reacts with hexokinase to give fructose 6-phosphate, as does another monosaccharide, *mannose*, which gives mannose 6-phosphate and then fructose 6-phosphate by way of the enzyme *phosphomannose isomerase* (Fig. 11-15).

Not all monosaccharides have such a simple or direct entry into the glycolytic pathway. Galactose, for example, is phosphorylated by *galactokinase* to galactose 1-phosphate, which then reacts with UTP via *galactose-1-phosphate uridylyltransferase* to yield UDP-galactose. This is converted into UDP-glucose by *UDP-glucose-4-epimerase*. The UDP-glucose can then be incorporated into glycogen and reappear as glucose 6-phosphate (Fig. 11-16).

While the entry of some monosaccharides (such as galactose) into the glycolytic pathway is circuitous, for others, alternative sequences of reactions are available. An alternative pathway exists for the metabolism of fructose, apart from its direct conversion into fructose 6-phosphate via hexokinase. This involves the conversion of fructose by way of *fructokinase* into fructose 1-phosphate,

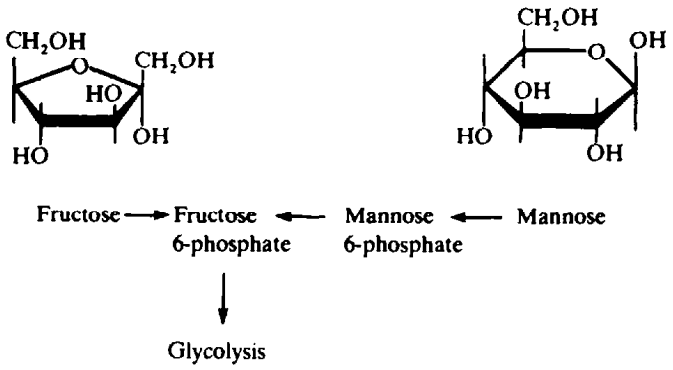


Fig. 11-15 The conversion of the monosaccharides fructose and mannose into the glycolytic intermediate fructose 6-phosphate.

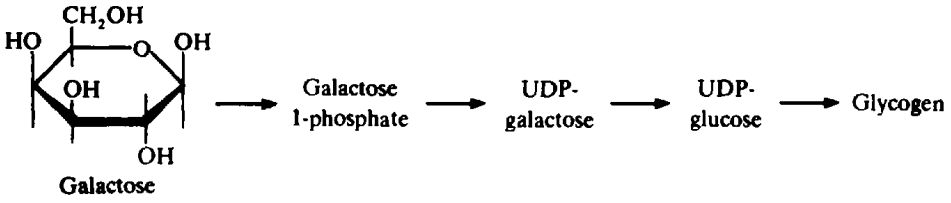


Fig. 11-16 The conversion of the monosaccharide galactose into glycogen.

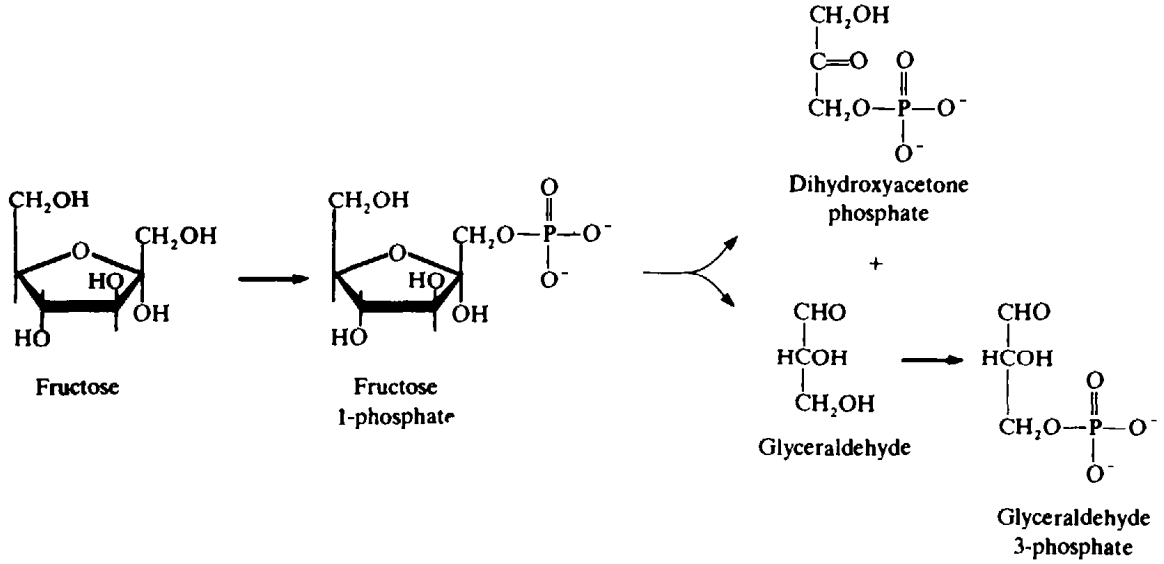


Fig. 11-17 The conversion of the monosaccharide fructose into the glycolytic intermediates dihydroxyacetone phosphate and glyceraldehyde 3-phosphate.

which is then cleaved by the action of *fructose 1-phosphate aldolase* into dihydroxyacetone phosphate and glyceraldehyde. Glyceraldehyde is then phosphorylated via *glyceraldehyde kinase* to give glyceraldehyde 3-phosphate. Fructose is, therefore, converted into dihydroxyacetone phosphate and glyceraldehyde 3-phosphate, two C_3 intermediates of the glycolytic pathway (Fig. 11-17).

Another C_3 compound that can enter the glycolytic pathway is *glycerol*. Glycerol is converted via *glycerol kinase* to *glycerol-3-phosphate*, which, with glycerol-3-phosphate dehydrogenase, produces dihydroxyacetone phosphate (Fig. 11-18).

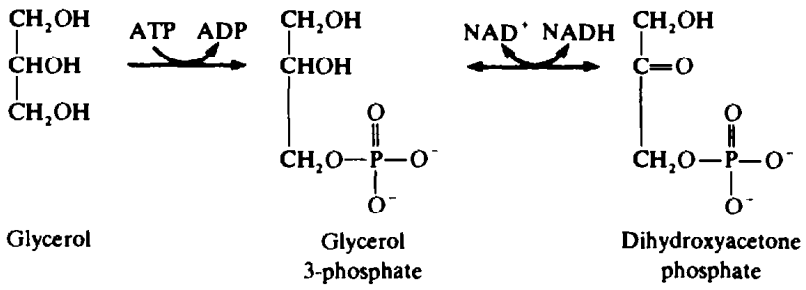


Fig. 11-18 The conversion of glycerol into the glycolytic intermediate dihydroxyacetone phosphate.

Question: What is the metabolic origin of glycerol?

It is derived principally from the hydrolysis of triglycerides (Chap. 13).

With one exception all these carbohydrates yield the same amount of ATP when they are degraded to pyruvate; that is, one mole of ATP per mole of pyruvate formed. The exception is galactose, which requires the equivalent of three molecules of ATP to produce glucose 1-phosphate and a further molecule of ATP to give fructose 1,6-bisphosphate, so the yield of ATP per molecule of pyruvate formed from galactose is 0.5.

EXAMPLE 11.5

One of the most abundant forms of D-glucose is cellulose. Can this be converted into a glycolytic intermediate in mammals?

No. Cellulose is a polysaccharide of D-glucose units joined by $\beta(1 \rightarrow 4)$ linkages, and mammalian cells cannot cleave this particular link because they do not possess the enzyme *cellulase*.

11.7 REGENERATION OF CYTOPLASMIC NAD^+ LEVELS

Most cells have the capacity to derive energy from three types of fuels: carbohydrates, amino acids, and fatty acids. By far the major source of energy is the fatty acids, which are degraded in the mitochondria to acetyl-CoA. Acetyl-CoA, which can also be formed from pyruvate, enters the citric acid cycle (Chap. 12), where it can be oxidized completely to CO_2. This process generates NADH and $FADH_2$, which are oxidized in the electron-transport system. Cells which do not possess mitochondria (e.g., red blood cells) or which contain very few (e.g., skeletal muscle), either are unable or have a diminished ability to use fatty acids as fuels. Thus, these cells rely on glycolysis as their means of phosphorylating ADP. To ensure that sufficient NAD^+, so vital to Step 6 in glycolysis, is always available, these cells contain mechanisms to regenerate cytoplasmic $NAD^+$from NADH; these mechanisms result in the conversion of pyruvate to lactate in mammalian cells (lactate fermentation; Sec. 11.2).

EXAMPLE 11.6

What mechanisms do nonmammalian cells use to regenerate NAD^+ for continued glycolysis?

Yeast, a facultative anaerobe, uses alcoholic fermentation (Sec. 11.2); pyruvate decarboxylase catalyzes the conversion of pyruvate to acetaldehyde, and then alcohol dehydrogenase converts the acetaldehyde to ethyl alcohol and oxidizes NADH to NAD^+.

Question: Cells that possess mitochondria can use pyruvate when oxygen is available. How do these cells cope with the problem of regenerating NAD^+?

The fact that mitochondria are present means that these cells can convert cytoplasmic NADH into NAD^+. The H's, or reducing equivalents, are transported into the mitochondria via metabolite *shuttles*.

Although mitochondria contain both NAD^+ and NADH, as does the cytoplasm, the mitochondrial and cytoplasmic pools are unable to exchange their contents directly, as the mitochondrial membranes are impermeable to the cytoplasmic compounds. The shuttle mechanisms allow the H on cytoplasmic NADH to be transported on other compounds into the mitochondria, where it is donated to NAD^+ (to form mitochondrial NADH) or to FAD (to form mitochondrial $FADH_2$). There are several shuttle mechanisms that are used by mammalian cells; two of the most important are the *malate-aspartate shuttle* and the *glycerol 3-phosphate shuttle*.

The Malate-Aspartate Shuttle

Cytoplasmic malate dehydrogenase reduces oxaloacetate to malate, thereby oxidizing cytoplasmic NADH to NAD^+ (see Example 11.1). The malate enters the mitochondria via a specific protein carrier (a dicarboxylate carrier) in the inner membrane and is oxidized to oxaloacetate by *mitochondrial* malate dehydrogenase, a process that is coupled to the reduction of mitochondrial NAD^+ to NADH (see Fig. 11-19). While this results in the shuttling of cytoplasmic H into the mitochondria, the cytoplasmic oxaloacetate must be replaced so that the process can continue. This is accomplished by a *transaminase* enzyme, which concomitantly converts oxaloacetate to *aspartate* and glutamate to *2-oxoglutarate* (Chap. 12). The aspartate and 2-oxoglutarate are transported via specific carriers into the cytoplasm, where another transaminase converts them back to oxaloacetate and glutamate, respectively.

The Glycerol 3-Phosphate Shuttle

Cytoplasmic *glycerol-3-phosphate dehydrogenase* reduces dihydroxyacetone phosphate to *glycerol 3-phosphate*, with the accompanying oxidation of NADH to NAD^+. The glycerol 3-phosphate passes through the outer mitochondrial membrane via a specific transport protein and is reoxidized to dihydroxyacetone phosphate by way of *mitochondrial* glycerol 3-phosphate dehydrogenase that is

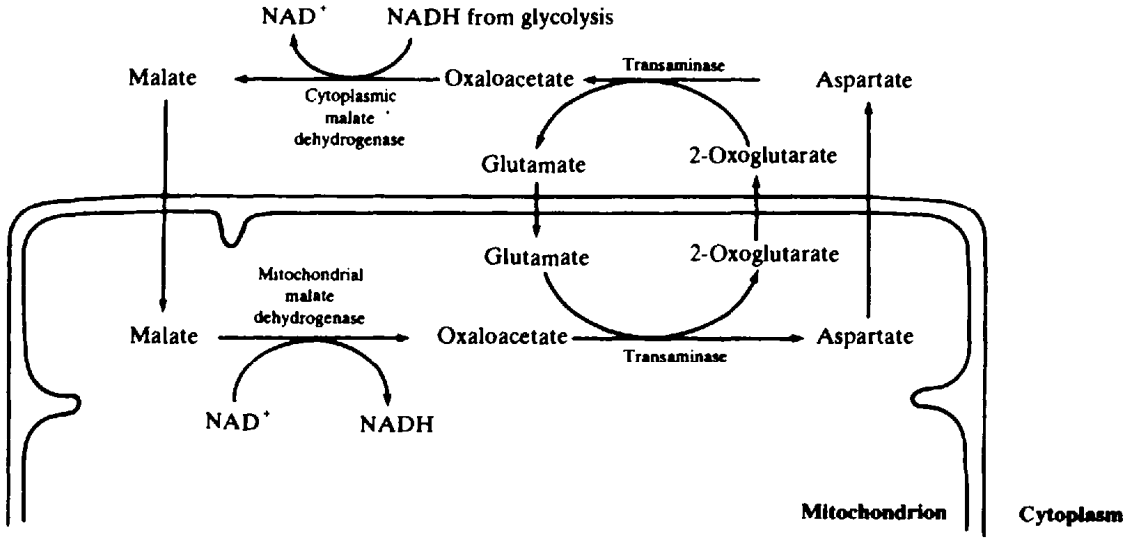


Fig. 11-19 The malate-aspartate shuttle, a mechanism for the transfer of reducing equivalents (H's) between cytoplasm and mitochondria.

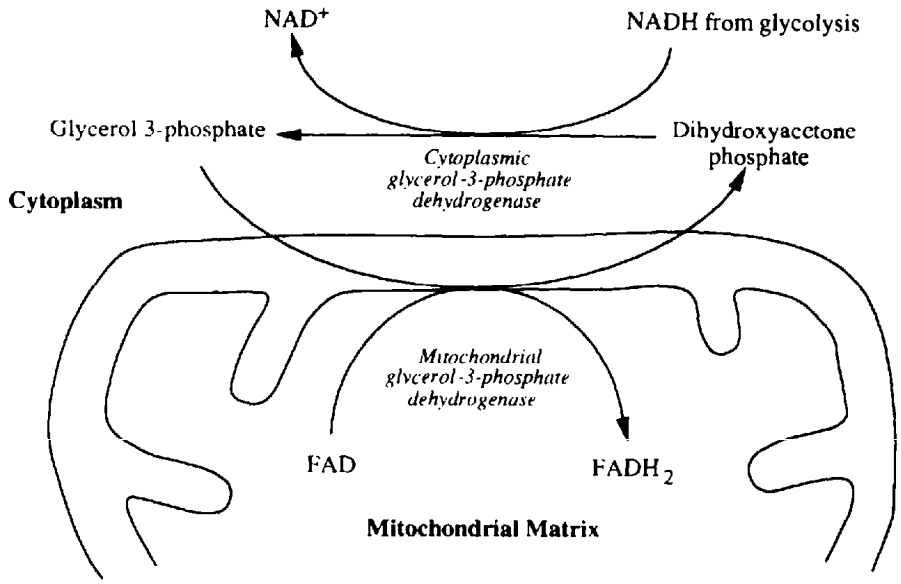


Fig. 11-20 The glycerol 3-phosphate shuttle, a mechanism for the transferral of reducing equivalents (H's) between the cytoplasm and mitochondria.

located within the inner mitochondrial membrane (Fig. 11-20). The redox coreactant for this enzyme is not NAD^+, but FAD (see Example 11.1, Fig. 11-2), and this is reduced to $FADH_2$. The dihydroxyacetone phosphate diffuses into the cytoplasm, where it is available for the shuttle to continue.

The net outcome of the two mechanisms just described is that cytoplasmic NADH is oxidized to cytoplasmic NAD^+. The H's, or reducing equivalents, are transferred to the mitochondria as NADH or $FADH_2$, according to the particular shuttle that is used.

11.8 CONTROL OF GLYCOLYSIS

Glycolysis has three irreversible steps, and the enzymes catalyzing these reactions exert control over the whole pathway. The major controlling enzyme is phosphofructokinase (Step 3). It is allosteric, is activated by ADP and AMP, and is inhibited by ATP, so that it is *most active* under conditions of low cellular energy and *least active* when the energy status of the cell is high. The activity of phosphofructokinase is also inhibited by NADH, which implies that if the NADH formed in glycolysis accumulates, the enzyme becomes inhibited until the NADH is oxidized to NAD^+. Further control by citrate and long-chain fatty acids suggests that when these compounds are abundant, the degradation of glucose is not essential, so that the reactions of the glycolytic pathway become inhibited. When phosphofructokinase is inhibited, fructose 6-phosphate and glucose 6-phosphate accumulate, and the latter substrate then exerts its control by *inhibiting* the activity of hexokinase (Step 1). This could be viewed as a second means of glycolytic control effected by phosphofructokinase. Pyruvate kinase (Step 10) is inactivated by high concentrations of ATP, citrate, and long-chain fatty acids and so acts in concert with phosphofructokinase.

Another mode of control is via glyceraldehyde-3-phosphate dehydrogenase (Step 6); this enzyme is activated by its oxidized coreactant (NAD^+) and inhibited by its reduced coreactant (NADH). Again, if the energy level of the cell is high, all the controlling enzymes cooperate to reduce the rate of glycolysis and, consequently, to conserve glucose.

The rate of glycolysis changes in a particular cell from moment to moment as its energy

requirements change. This rate is referred to as the *glycolytic flux*. If, for example, a cell requires energy and can obtain this only from glucose, then the glycolytic flux would be high. But if a cell has sufficient energy for its immediate requirements as a result of degrading glucose or any other energy-producing compound, then the controls would operate and the glycolytic flux would be low.

EXAMPLE 11.7

A good illustration of how glycolytic flux can rapidly change is seen when yeast are grown under aerobic and anaerobic conditions with glucose as the carbon source. This effect, first observed by Louis Pasteur, is called the *Pasteur effect* and is depicted in Fig. 11-21.

When grown under *aerobic* conditions, the yeast produces two ATP molecules from one molecule of glucose by substrate-level phosphorylation in glycolysis. The two molecules of pyruvate produced can then be completely oxidized to CO_2, and each yields a further 15 molecules of ATP. This leads to a slow decrease in the concentration of glucose, a steady production of CO_2, and relatively little change in the amount of ATP. Also, the two molecules of NADH can be reoxidized to NAD^+ by the electron-transport system. (This produces yet more ATP, as discussed in Chap. 14.)

When yeast are grown in an *anaerobic* environment, the utilization of glucose is markedly increased and the production of CO_2 and ethanol rises dramatically with very little change in the concentration of ATP. While this phenomenon has been explained in many ways, the simplest requires the assumption that the yeast needs a constant amount of ATP for its energy requirements, irrespective of the conditions under which it is grown.

When yeast are deprived of oxygen, the generation of ATP is possible only by using the reactions of the glycolytic pathway. Furthermore, the ability of the yeast to regenerate NAD^+ by the electron-transport system is denied in the absence of oxygen. If this situation is not modified, the level of ATP will decline and the cells will head toward a low-energy state. Controls in glycolysis are now lifted and the glycolytic flux increases. To cope with this, alcoholic fermentation comes into operation to remove the pyruvate as ethanol and to regenerate NAD^+ so that glycolysis can continue. As this permits only two molecules of ATP to be generated per molecule of glucose consumed, the rate of glycolysis increases; ethanol is produced and glucose is consumed more quickly than under aerobic conditions. The production of CO_2 also increases even though the maximum theoretical yield is two molecules per molecule of glucose (anaerobic) as against a possible six molecules per molecule of glucose under aerobic conditions. So, to maintain a constant concentration of ATP, yeast grown in an aerobic environment require the consumption of less glucose than do those grown under anaerobic conditions. Also, aerobic growth will result in the production of less CO_2 than will growth under anaerobic conditions.

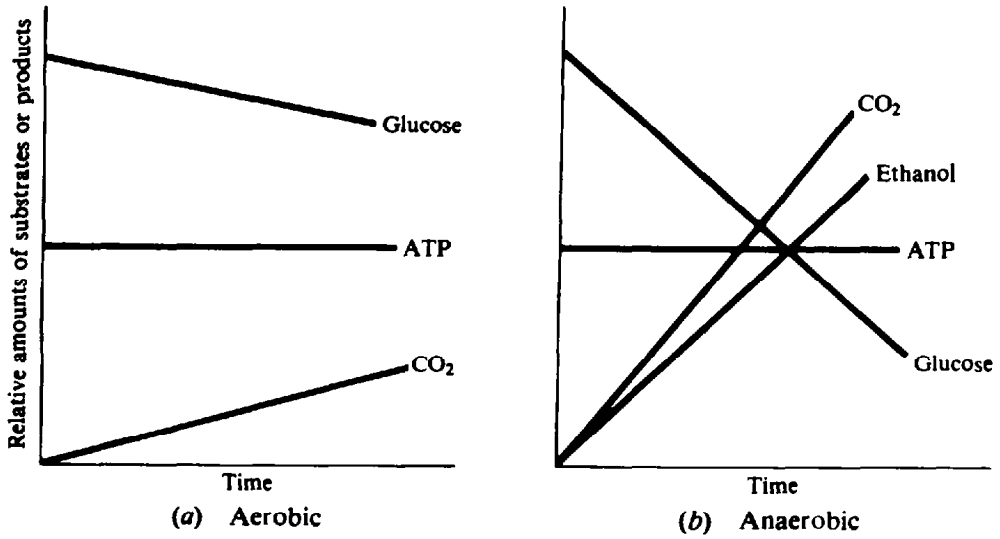


Fig. 11-21 The Pasteur effect: metabolic activity of yeast grown under aerobic and anaerobic conditions.

11.9 EFFECTS OF HORMONES ON GLYCOLYSIS

Hormones do not exert any direct control on the rate of glycolysis, but three hormones have some indirect influence. These three are *insulin*, *glucagon*, and *epinephrine*. Basically, insulin is involved in the transport of glucose into all cells (except liver and red blood cells), whereas glucagon and epinephrine are both concerned with the degradation of glycogen in the liver; epinephrine is also involved with the degradation of glycogen in muscle.

EXAMPLE 11.8

Where are insulin and glucagon produced, and what controls their secretion into the blood?

Insulin and glucagon are polypeptide hormones synthesized in, and secreted by, the *pancreas*. Insulin is produced by the β cells of the pancreas, and glucagon by the α cells. The secretion of either of these hormones depends on the blood glucose concentration: above 4.5–5.5 m*M* (80–100 mg/100 mL) of glucose, insulin is secreted, but below 4.5 m*M* (80 mg/100 mL), glucagon is secreted.

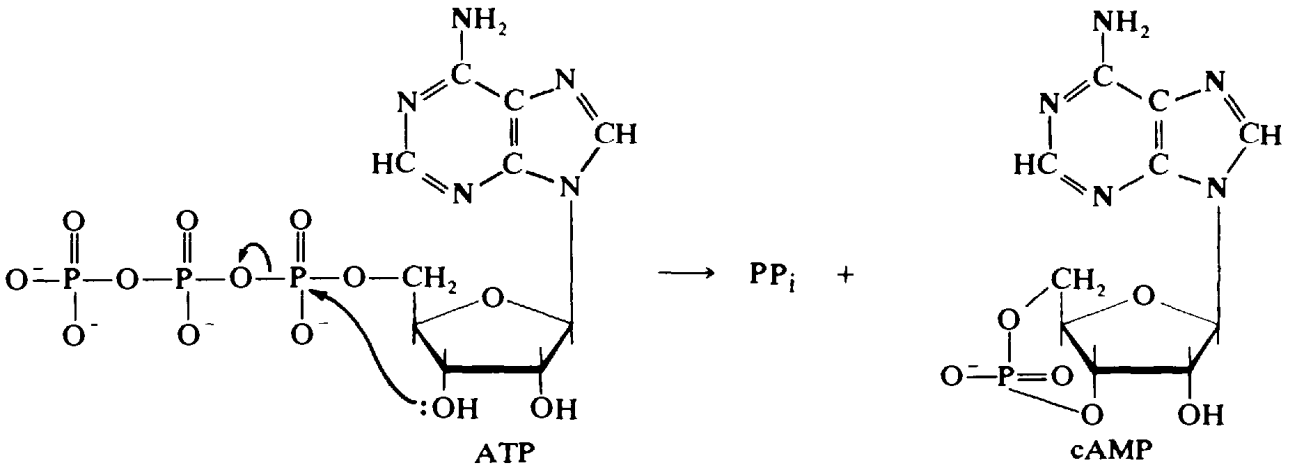
Insulin and glucagon have opposing effects; insulin *reduces* blood glucose concentration by several means, including (1) the enhancement of glucose uptake into muscle and fat cells, (2) the activation of glycogen synthase, and (3) the stimulation of the activities of phosphofructokinase and glucokinase. The overall effect of insulin is that blood glucose concentration *decreases* and the glucose is converted into glycogen, used in glycolysis, or both. The action of glucagon is one of *increasing* blood glucose concentration by stimulating the degradation of liver glycogen to glucose 1-phosphate, which can be converted to glucose 6-phosphate by phosphoglucomutase and then to glucose by glucose 6-phosphatase.

Epinephrine is synthesized within the *adrenal cortex* and is rapidly released into the bloodstream in response to a sudden fright or shock, which may signal the necessity for rapid evasive action requiring instantaneous muscular activity (fight or flight reflex). This hormone also stimulates the degradation of liver glycogen, as well as the degradation of muscle glycogen to glucose 6-phosphate in anticipation of its requirement in muscle glycolysis.

The action of glucagon and epinephrine in the stimulation of glycogen degradation in their target tissues is mediated by a process involving *cyclic AMP* (Chap. 6).

EXAMPLE 11.9

Cyclic AMP, or *cAMP* for short, is produced by the intramolecular cyclization of ATP, a reaction catalyzed by the enzyme *adenylate* (or *adenyl*) *cyclase*.



Being hormones derived from many amino acids (glucagon) or from a single amino acid (epinephrine), these compounds cannot enter their target cells directly. Consequently, liver cells have separate

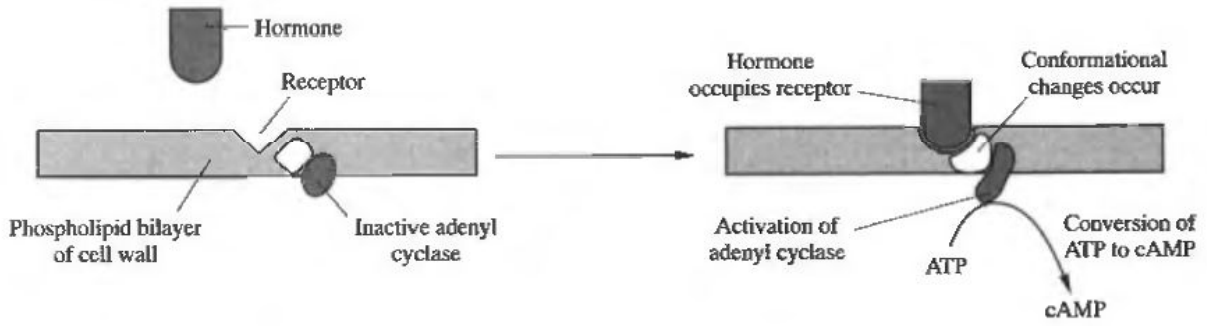


Fig. 11-22 The activation of adenyl cyclase by glucagon or epinephrine.

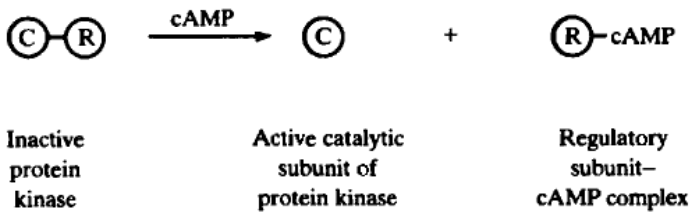


Fig. 11-23 The activation of protein kinase by cAMP.

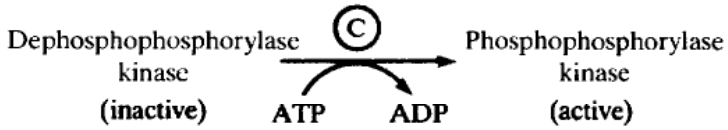


Fig. 11-24 The activation of phosphorylase kinase by protein kinase.

receptors on their outer membrane for glucagon and epinephrine, and muscle cells have only the receptor for epinephrine. When occupied by the hormone(s), these receptors undergo a conformational change. This change is transmitted through the membrane to the inner surface and results in the activation of adenyl cyclase. This enzyme, which lies under the receptors, is dormant or inactive when the receptors are unoccupied but is activated when the receptors are occupied. Once activated, adenyl cyclase catalyzes the cyclization of ATP to cAMP (Fig. 11-22).

The cAMP is called the *second messenger*. It is produced within the liver or muscle cell when the specific hormone (the *first messenger*) occupies its particular receptor. The effect of cAMP is the activation of an enzyme, a *protein kinase*, by binding to and removing the enzyme's regulatory subunit (Fig. 11-23).

The active protein kinase then uses ATP to activate another enzyme, called *phosphorylase kinase*, by phosphorylating it (Fig. 11-24).

Finally, the active phosphorylase kinase converts the inactive form of another enzyme, *phosphorylase* b, into its active form, *phosphorylase* a (Fig. 11-25).

Phosphorylase *a* is the active form of glycogen phosphorylase, and this can now catalyze the conversion of glycogen to glucose 1-phosphate. The entire process is shown in Fig. 11-26, which depicts the action of epinephrine on a liver or muscle cell.

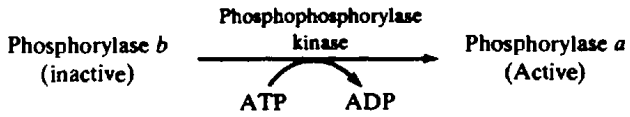


Fig. 11-25 The conversion of phosphorylase *b* to phosphorylase *a* via phosphorylase kinase.

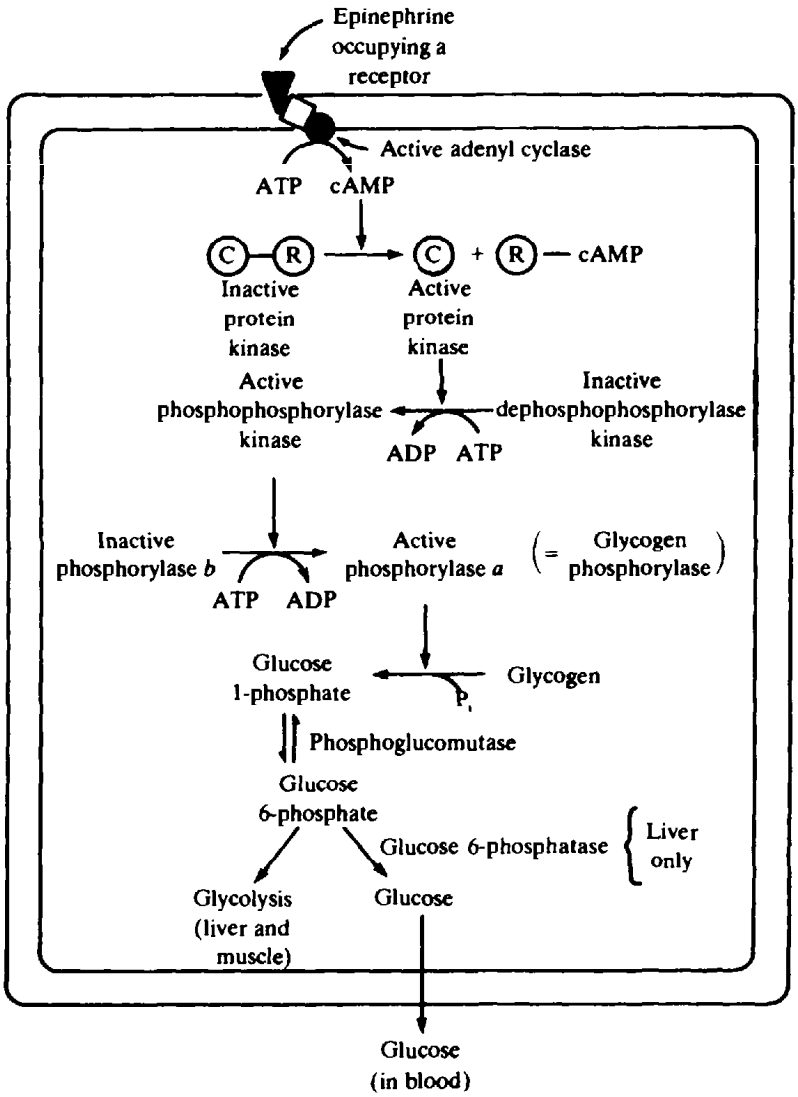


Fig. 11-26 The action of epinephrine on the degradation of glycogen in a liver or muscle cell.

Question: What is the purpose of this multistep *cascade* and what stops it?

Each step requires activation of an enzyme, and each enzyme can perform the next step many times. This gives a *cascade* effect, in which each step is an *amplification* of the one preceding it. There are a number of ways in which the process can slow down or stop. The epinephrine can leave its receptor; this would cause adenyl cyclase to assume its inactive or dormant conformation. The cAMP can be hydrolyzed to AMP; this process is catalyzed by the enzyme *phosphodiesterase*. The concentration of ATP will increase (as the result of the degradation of glycogen and the production

of ATP by substrate-level phosphorylation in glycolysis); this will allosterically inhibit the conversion of phosphorylase *b* to phosphorylase *a*.

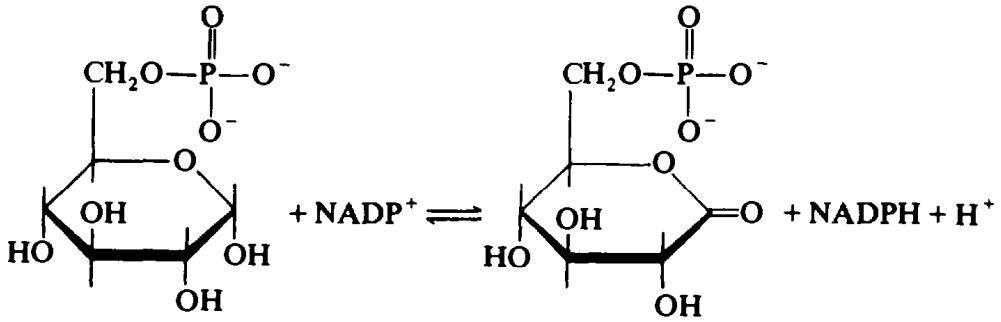
Under normal conditions the activity of glycogen phosphorylase (phosphorylase *a*) is regulated by the cellular concentration of glucose 6-phosphate, ATP, and AMP; the first two effectors inhibit its activation, and the third promotes its activation.

The mechanism of action of glucagon on the mobilization of glycogen in the liver is similar to that of epinephrine. The function of glucagon is to increase the concentration of blood glucose to normal levels, which is the opposing effect to that of insulin. Epinephrine has an immediate action that leads to the rapid degradation of glycogen in muscle and liver cells, so that muscles have an abundance of glucose 6-phosphate for glycolysis coupled with a plentiful supply of glucose available in the blood as a result of the ability of the liver to convert its glycogen rapidly to glucose.

11.10 THE PENTOSE PHOSPHATE PATHWAY

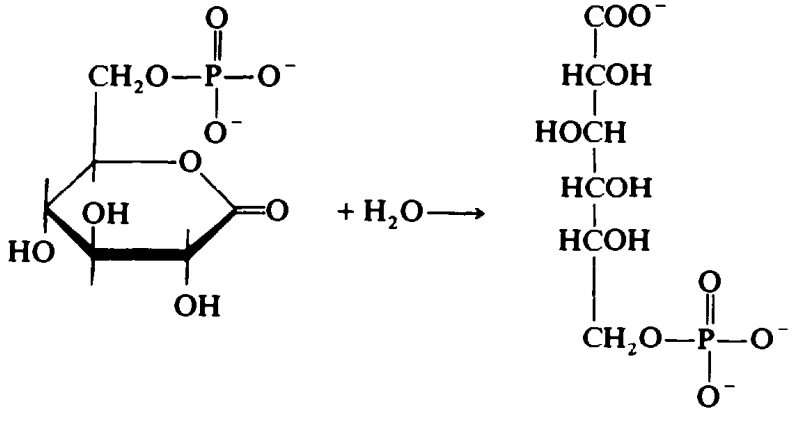
Some mammalian cells have the ability to metabolize glucose 6-phosphate in a pathway that involves the production of C_3, C_4, C_5, C_6, and C_7 sugars. This process also yields the reduced coenzyme, NADPH, which is oxidized in the biosynthesis of fatty acids and steroids (Chap. 13). Consequently, this metabolic pathway is of major importance in those cells involved in fatty acid and steroid production, such as the liver, lactating mammary gland, adrenal cortex, and adipose tissue. The pentose phosphate pathway, which does not require oxygen and which occurs in the cytoplasm of these cells, has two other names: the *phosphogluconate pathway* (after the first product in the pathway) and the *hexose monophosphate shunt* (since the end products of the pathway can reenter glycolysis).

The first step is the oxidation of glucose 6-phosphate at C-1, catalyzed by glucose-6-phosphate dehydrogenase, to produce a cyclic ester (also called a *lactone*):



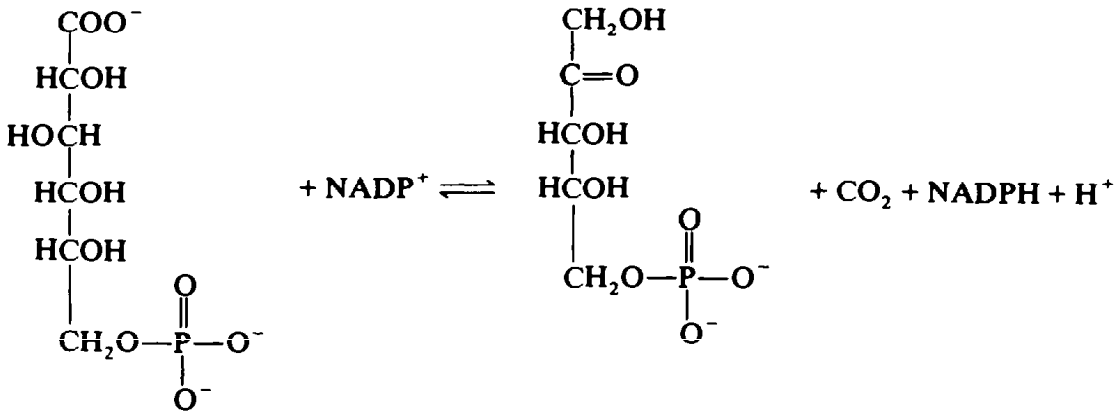
Glucose 6-phosphate — 6-Phosphogluconolactone

The 6-phosphogluconolactone is unstable, and the ester spontaneously hydrolyzes, although a specific enzyme (lactonase) also catalyzes this reaction (Step 2):



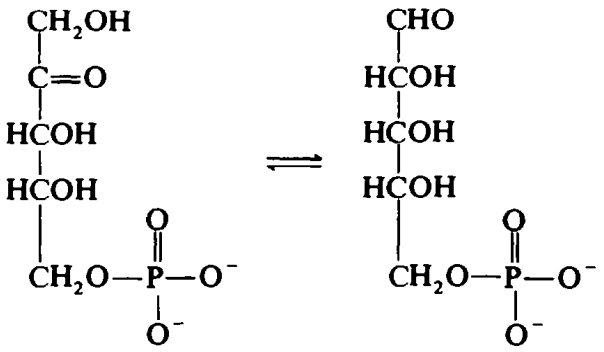
6-Phosphogluconolactone — 6-Phosphogluconate

The third step in the pathway yields a C_5 sugar, D-ribulose 5-phosphate, and concomitantly reduces another molecule of $NADP^+$. This is catalyzed by *6-phosphogluconate dehydrogenase*:



6-Phosphogluconate D-Ribulose 5-phosphate

D-Ribulose 5-phosphate then undergoes an isomerization by *ribose-5-phosphate isomerase* to D-ribose 5-phosphate (Step 4):



D-Ribulose 5-phosphate D-Ribose 5-phosphate

These four steps constitute the *first phase* of the pentose phosphate pathway and result in the consumption of one molecule of glucose 6-phosphate and the formation of one molecule of ribose 5-phosphate, two of NADPH, and one of CO_2.

The requirement for NADPH far exceeds an equal requirement for ribose 5-phosphate (necessary for the production of nucleic acids and nucleotides), and so the *second phase* of the pentose phosphate pathway converts the C_5 sugar, by a series of reversible reactions, into the glycolytic intermediates fructose 6-phosphate and glyceraldehyde 3-phosphate. This interconversion is shown in Fig. 11-27. Not only does the second phase of the pathway conserve all the carbon atoms of the C_5 sugar, but it produces erythrose 4-phosphate (C_4), xylulose 5-phosphate (C_5), and sedoheptulose 7-phosphate (C_7), which are available to other metabolic processes.

EXAMPLE 11.10

What is the overall reaction of the pentose phosphate pathway?
The balanced sequence is

$$3 \text{ Glucose 6-phosphate} + 6NADP^+ \rightarrow 3CO_2 + 6NADPH + 2 \text{ Fructose 6-phosphate} + \text{Glyceraldehyde 3-phosphate}$$

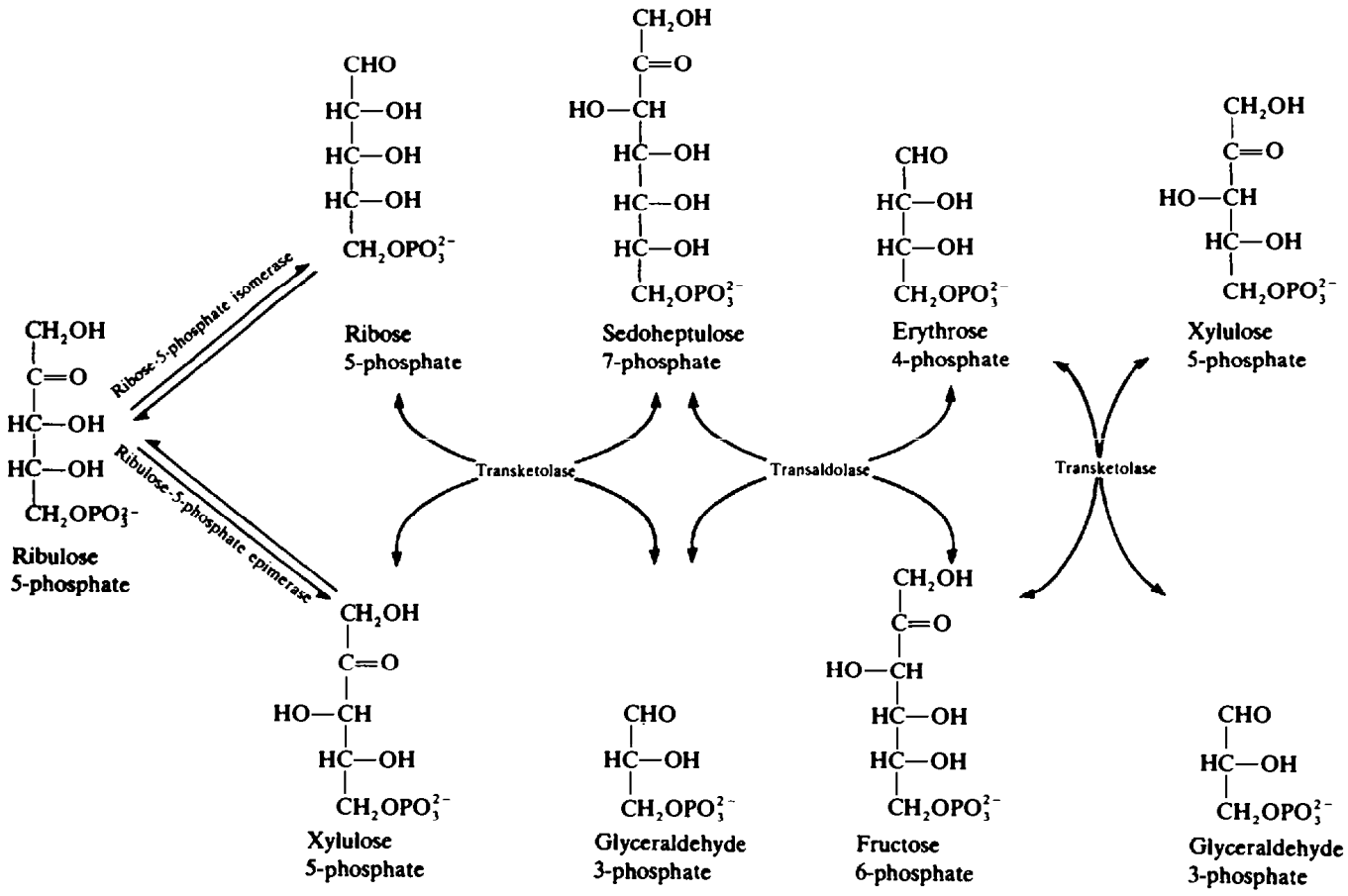


Fig. 11-27 The second phase of the pentose phosphate pathway.

This shows the production of NADPH in the pathway and its return of all but one of the carbon atoms of glucose 6-phosphate to the glycolytic pathway. However, the overall scheme belies its adaptability to produce and to deal with a variety of sugar phosphates.

Solved Problems

GLYCOLYSIS

11.1. A compound is an inhibitor of glyceraldehyde-3-phosphate dehydrogenase. If this compound were added to liver cells where D-glucose was the only substrate, what effect would it have on the concentrations of the glycolytic intermediates?

SOLUTION

There would be an accumulation of those intermediates from glucose 6-phosphate to glyceraldehyde 3-phosphate and a depletion of those from 1,3-bisphosphoglycerate to pyruvate.

11.2. If the substrate for the liver cells in Prob. 11.1 were L-lactate, what effect would the inhibitor have on the concentrations of the glycolytic intermediates?

SOLUTION

There would be no effect on the concentrations of the glycolytic intermediates except, perhaps, an increase in the concentration of pyruvate. The cells would convert the lactate to pyruvate and use this as a precursor of acetyl-CoA in the citric acid cycle. If glucose or any other suitable carbohydrate is not available to the cell, then glycolysis cannot operate.

11.3. What constraint prevents the intermediates of the glycolytic pathway from leaving the cell in which they are formed?

SOLUTION

All the glycolytic intermediates are phosphorylated. At physiological pH, these phosphate groups are ionized so that each intermediate is negatively charged. Charged molecules are not readily able to cross membranes (unless there is a specific transport protein), and the intermediates are, therefore, confined to the cytoplasm of the cell.

11.4. What is the net oxidation change when glucose is converted to lactate?

SOLUTION

There is no overall change in the oxidation state when glucose is converted to lactate, because glyceraldehyde-3-phosphate dehydrogenase *oxidizes* glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate, but lactate dehydrogenase *reduces* pyruvate to lactate. These two reactions also *reduce* NAD^+ to NADH, then reoxidize the NADH to NAD^+.

11.5. If all the glycolytic enzymes, ATP, ADP, NAD^+, and glucose were incubated together under ideal conditions, would pyruvate be produced?

SOLUTION

No, because an important omission is inorganic phosphate. Even if phosphate were added to the incubation mixture, pyruvate would be produced only in an amount equivalent to that of the NAD^+ present. Glycolysis requires NAD^+ for Step 6, the reaction catalyzed by glyceraldehyde-3-phosphate dehydrogenase.

THE FATE OF PYRUVATE

11.6. What is the metabolic fate of lactate in mammalian cells?

SOLUTION

Lactate can undergo only one reaction: it is oxidized to pyruvate via lactate dehydrogenase.

THE CORI CYCLE

11.7. From where do skeletal muscles obtain their carbohydrate for glycolysis?

SOLUTION

One source is blood glucose. Glucose entering a muscle cell is rapidly converted to glucose 6-phosphate, which enters glycolysis. Another source is muscle glycogen, which is degraded to glucose 1-phosphate. The latter is then converted to glucose 6-phosphate, which proceeds through glycolysis.

ENTRY OF OTHER CARBOHYDRATES INTO GLYCOLYSIS

11.8. Which of the following could theoretically yield the maximum net number of molecules of ATP by substrate-level phosphorylation in glycolysis: a molecule of sucrose, two molecules of glucose, or two molecules of fructose?

SOLUTION

All can yield four molecules of ATP. Sucrose is converted into one molecule each of glucose and fructose; each of these sugars requires two molecules of ATP to reach the stage of two molecules of glyceraldehyde 3-phosphate. From here to pyruvate, each glyceraldehyde 3-phosphate molecule yields two molecules of ATP by substrate-level phosphorylation of ADP. Thus, one molecule of glucose or of fructose generates two molecules of ATP.

11.9. Which reactions are catalyzed by the enzymes sucrase and lactase?

SOLUTION

Sucrase catalyzes the hydrolysis of sucrose to glucose and fructose. Lactase catalyzes the hydrolysis of lactose to glucose and galactose.

CONTROL OF GLYCOLYSIS

11.10. If a molecule of glucose produces two molecules of ATP by substrate-level phosphorylation of ADP in glycolysis and the resulting two molecules of pyruvate can each yield 15 molecules of ATP when oxygen is available, how many glucose molecules will be necessary to produce 160 molecules of ATP by yeast grown under (*a*) aerobic and (*b*) anaerobic conditions?

SOLUTION

(*a*) Growth under aerobic conditions can produce $2 + 30 = 32$ molecules of ATP per molecule of glucose. To produce 160 molecules of ATP, $160 \div 32 = 5$ molecules of glucose are required.

(*b*) Growth under anaerobic conditions can produce only two molecules of ATP per molecule of glucose. To produce 160 molecules of ATP, $160 \div 2 = 80$ molecules of glucose are required.

11.11. In Prob. 11.10, how many molecules of CO_2 would be evolved in producing 160 molecules of ATP during growth under (*a*) aerobic and (*b*) anaerobic conditions?

SOLUTION

(*a*) Aerobic:

$$1 \text{ Glucose} \rightarrow 6CO_2$$
$$\therefore 5 \text{ Glucose} \rightarrow 30CO_2$$

(*b*) Anaerobic:

$$\therefore 1 \text{ Glucose} \rightarrow 2CO_2$$
$$80 \text{ Glucose} \rightarrow 160CO_2$$

Supplementary Problems

11.12. In glycolysis, the conversion of dihydroxyacetone phosphate to glyceraldehyde 3-phosphate, catalyzed by triosephosphate isomerase, is reversible. Given that at equilibrium the reaction strongly favors the formation of dihydroxyacetone phosphate, how does glycolysis proceed?

11.13. In red blood cells, 2,3-bisphosphoglycerate decreases the affinity of oxygen for hemoglobin by stabilizing the hemoglobin in its deoxygenated form (Chap. 5). This effector is synthesized in one step from a glycolytic intermediate and is converted to another glycolytic intermediate also in one step. What are the two glycolytic intermediates, and which enzymes catalyze the two reactions?

11.14. What is the fate of the six carbon atoms of glucose when it is metabolized by yeast grown under anaerobic conditions?

11.15. Calculate the number of high-energy phosphate bonds that are required for the conversion of two molecules of pyruvate into a glucose unit within glycogen.

11.16. A liver cell has a high concentration of glucose 6-phosphate, which inhibits the activity of hexokinase. What is the fate of the glucose 6-phosphate?

11.17. Termites exist almost entirely on a diet of cellulose. How is the cellulose degraded by the termites to carbohydrates, which can be used as an energy source?

11.18. When yeast are incubated in an oxygenated medium containing all necessary nutrients, and with glucose as the sole carbon source, the concentration of glucose decreases with time, carbon dioxide is evolved, and the concentrations of ADP and ATP remain fairly constant. Explain the metabolic processes that occur when the incubation is performed (*a*) in the absence of oxygen or (*b*) in the presence of both oxygen and an inhibitor of glyceraldehyde-3-phosphate dehydrogenase.

11.19. How is the concentration of glucose in the blood maintained (*a*) during rest after a carbohydrate-rich meal and (*b*) during prolonged exercise?

11.20. Diabetes mellitus is a condition in which the pancreas fails to produce sufficient active insulin or produces an ineffective form of insulin. What effects on carbohydrate metabolism would this condition impart to an otherwise healthy person?

11.21. For a sequence of reactions a cell requires NADPH far in excess of its requirements for ribose 5-phosphate. (*a*) How can the cell achieve this, and (*b*) what is the fate of the excess ribose 5-phosphate?

11.22. What are the four reactions that occur in liver cells that involve glucose 6-phosphate as a substrate?

Chapter 12

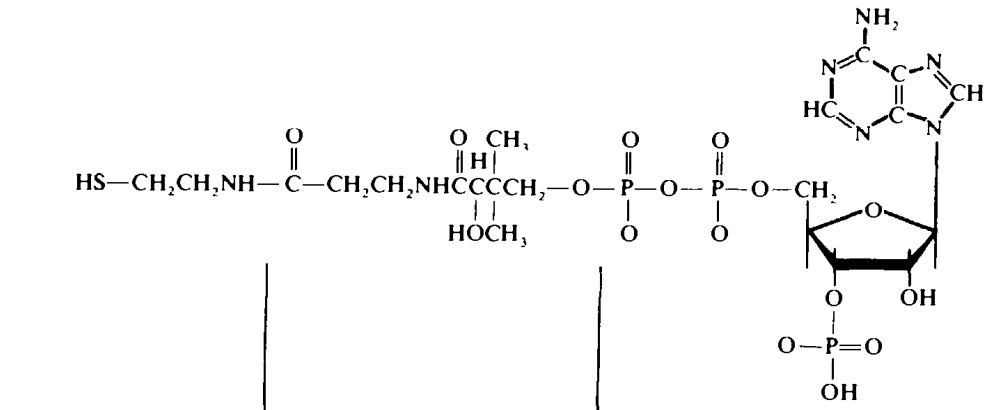
The Citric Acid Cycle

12.1 INTRODUCTION

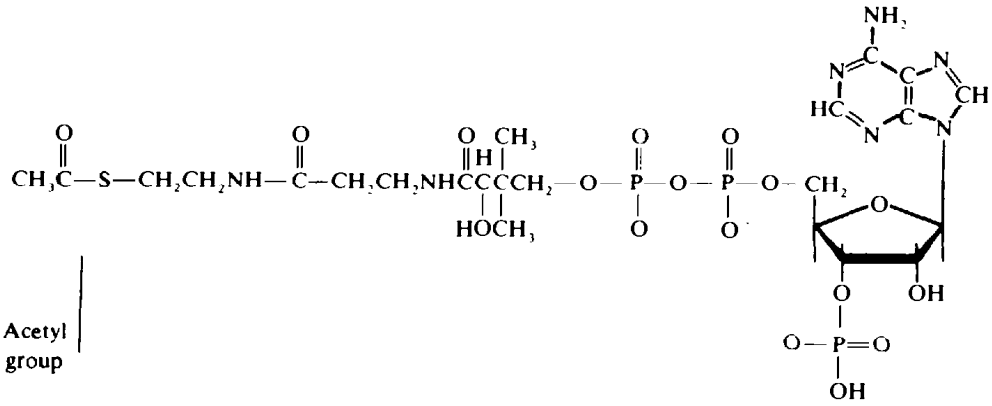
The citric acid cycle is a sequence of reactions in which the two carbon atoms of acetyl-CoA are ultimately oxidized to CO_2. It is the central pathway for the release of energy from acetyl-CoA, which is produced from the catabolism of carbohydrates (Chap. 11), fatty acids (Chap. 13), and some amino acids (Chap. 15) and is closely involved with two other processes, namely, electron transport and oxidative phosphorylation (Chap. 14).

Question: What is acetyl-CoA?

The name "acetyl-CoA" is an abbreviation for the compound acetyl coenzyme A, which has the structure shown in Fig. 12-1. Coenzyme A has three components: ADP with an additional 3'



Coenzyme A



Acetyl coenzyme A

Fig. 12-1 The structures of coenzyme A and acetyl coenzyme A.

phosphate group, pantothenic acid, and β-mercaptoethylamine. Coenzyme A is a *carrier* of *acyl* groups, especially acetyl groups, which are attached to the thiol of β-mercaptoethylamine. Thus, the abbreviation CoA is used for coenzyme A and acetyl-CoA for acetyl coenzyme A. In some cases the thiol group is emphasized by incorporating SH into the abbreviations; i.e., CoASH (coenzyme A) and $CoASCOCH_3$ or $CH_3COSCoA$ (acetyl coenzyme A).

EXAMPLE 12.1

The degradation of carbohydrates (such as glucose) in the glycolytic pathway produces pyruvate, but how does the acetyl-CoA originate in carbohydrate metabolism?

Pyruvate is converted into acetyl-CoA by a group of enzymes known as the *pyruvate dehydrogenase complex* (see Example 12.3 and Chap. 5). Acetyl-CoA and the enzymes that catalyze the steps of the citric acid cycle are situated within the matrix of the mitochondria, except for one enzyme that is located in the inner mitochondrial membrane.

Question: Why is the citric acid cycle so named?

The name stems from the first step in the cycle, which is a condensation of oxaloacetate with acetyl-CoA to form *citric acid*. However, as this product is a tricarboxylic acid, the cycle has an alternative name, the *tricarboxylic acid cycle*.

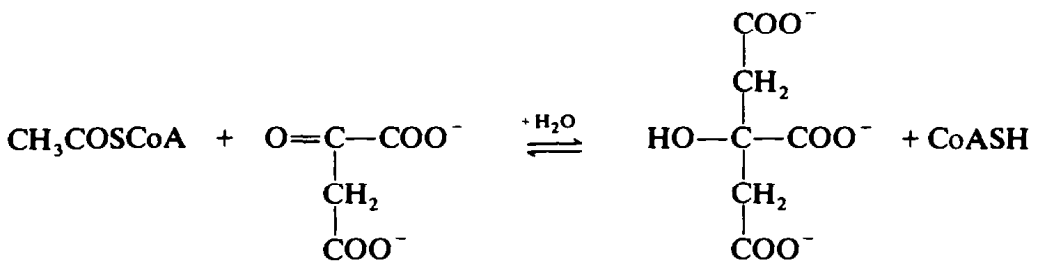
Acetyl-CoA cannot permeate the two mitochondrial membranes, and it is formed exclusively in the mitochondrial matrix from three basic sources:

1. Glycolysis produces pyruvate, which readily crosses the mitochondrial membranes via a transport protein and enters the matrix, where it is converted to acetyl-CoA via pyruvate dehydrogenase.
2. Fatty acids enter from the cytoplasm (as their CoA derivatives) via a specific transport protein, and in the matrix, they undergo oxidation to acetyl-CoA.
3. Proteins are hydrolyzed to amino acids, and those amino acids that are converted into acetyl-CoA (Chap. 15) undergo this reaction within the mitochondrial matrix.

12.2 REACTIONS OF THE CITRIC ACID CYCLE

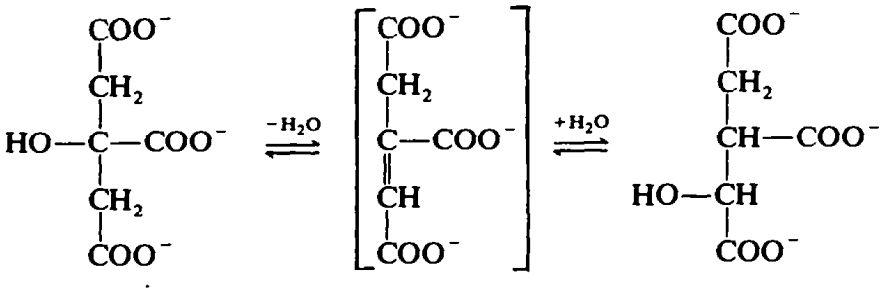
There are eight steps in the citric acid cycle.

Step 1 is the condensation of acetyl-CoA with oxaloacetate to form citrate, catalyzed by *citrate synthase*.



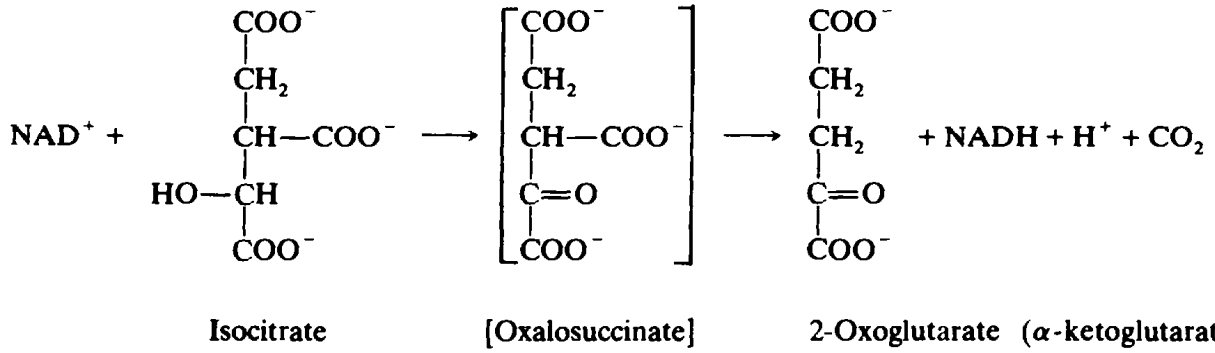
Acetyl-CoA Oxaloacetate Citrate Coenzyme A

Step 2 involves the dehydration of citrate to *cis*-aconitate followed by the hydration of *cis*-aconitate to isocitrate. *Aconitase* catalyzes these reversible reactions.



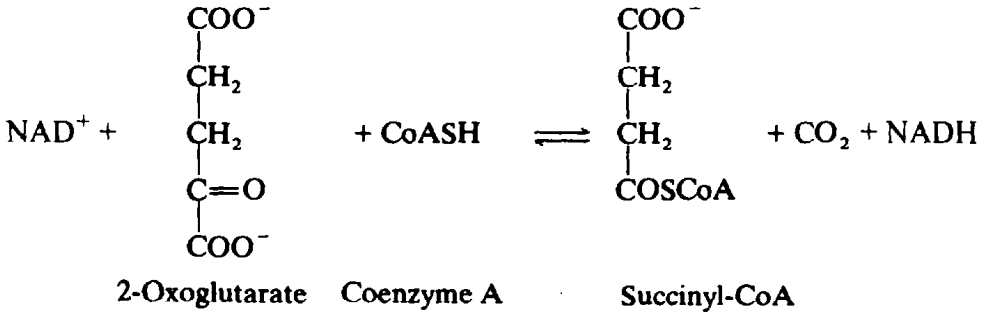
The intermediate in this reaction, *cis*-aconitate, is bound to aconitase and is not usually classed as a discrete intermediate of the citric acid cycle.

Step 3 involves *isocitrate dehydrogenase* which oxidatively decarboxylates isocitrate to 2-oxoglutarate.

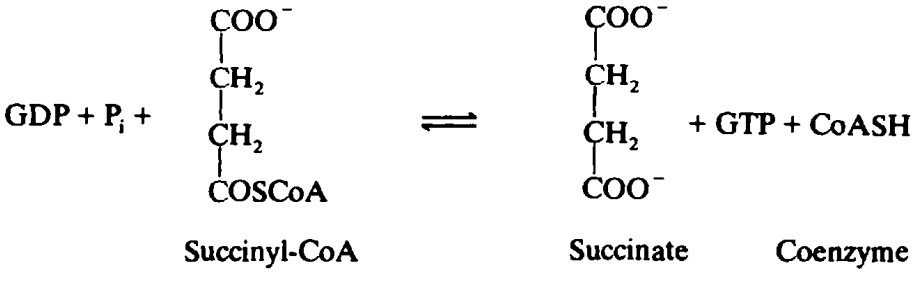


The intermediate in this reaction, *oxalosuccinate*, does not dissociate from the enzyme and is not usually classed as a discrete intermediate of the citric acid cycle.

Step 4 is the production of succinyl-CoA from 2-oxoglutarate and coenzyme A (CoASH), catalyzed by the *2-oxoglutarate dehydrogenase complex* (which is often called *α-ketoglutarate dehydrogenase*).



Step 5 is the conversion of succinyl-CoA to succinate, catalyzed by *succinyl-CoA synthetase*.

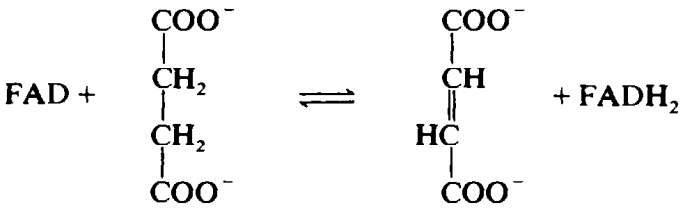


The phosphorylation of GDP in this reaction is an example of *substrate-level phosphorylation*, and this is the only reaction in the citric acid cycle to produce a high-energy phosphate bond directly. The energy for this phosphorylation is derived from the hydrolysis of the *thioester bond* of succinyl-CoA. Subsequently, GTP phosphorylates ADP, catalyzed by *nucleoside diphosphokinase*, but this reaction

$$GTP + ADP \rightleftharpoons GDP + ATP$$

is not metabolically an essential part of the citric acid cycle.

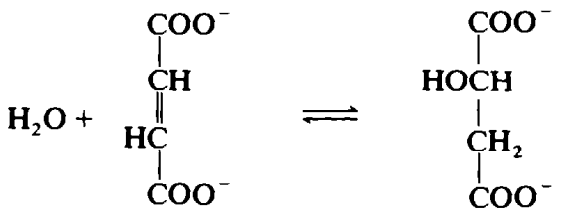
Step 6 is the oxidation of succinate to fumarate, catalyzed by *succinate dehydrogenase*.



Succinate Fumarate

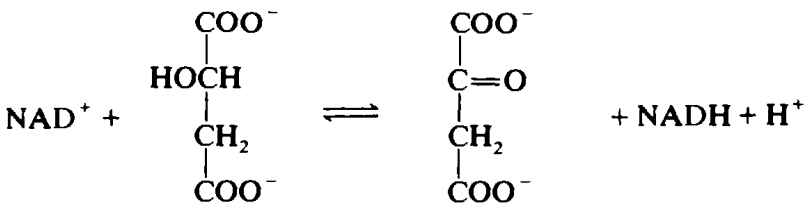
Succinate dehydrogenase catalyzes the so-called *trans elimination* of two H's. This is the only reaction in the citric acid cycle involving FAD, and succinate dehydrogenase is the only enzyme in the cycle that is *membrane-bound*. The importance of this will be discussed in Chap. 14.

Step 7 is the reversible hydration of fumarate to form malate, catalyzed by *fumarate hydratase* (which is usually called *fumarase*).



Fumarate L-Malate

Step 8, catalyzed by *malate dehydrogenase*, is the formation of oxaloacetate from malate.



L-Malate Oxaloacetate

With the regeneration of oxaloacetate in step 8, the cycle of reactions is completed, and the oxaloacetate condenses with another molecule of acetyl-CoA to commence another turn of the cycle. However, note that there is actually no beginning or end to the citric acid cycle (Fig. 12-2); if any of the intermediates are produced within the mitochondria or gain access to the mitochondria, they can participate in the cycle of reactions but they will always be regenerated. The only actual *fuel* for the cycle is acetyl-CoA and this is not regenerated; its carbon atoms are liberated at steps 3 and 4 as CO_2.

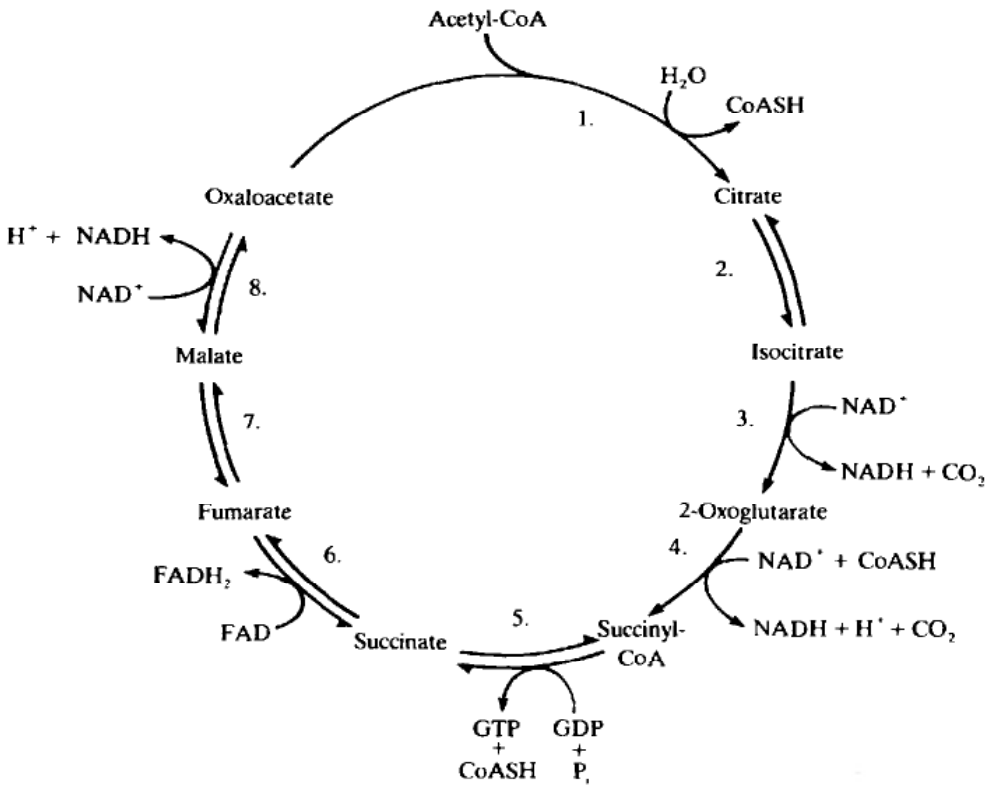


Fig. 12-2 The citric acid cycle. The numbered reactions refer to the steps given in the text.

Question: The citric acid cycle is part of the process called *respiration*. Is oxygen directly involved in the reactions of the cycle?

No. While oxygen is not involved in any of the steps of the citric acid cycle, CO_2 is liberated at two of them, giving rise to part of the stoichiometry of respiration.

12.3 THE ENERGETICS OF THE CITRIC ACID CYCLE

The overall consumption of one molecule of acetyl-CoA in the citric acid cycle is an *exergonic* process; $\Delta G^{\circ\prime} = -60\,\text{kJ mol}^{-1}$. All but two of the individual reactions are exergonic. Step 2 (citrate → isocitrate) and step 8 (malate → oxaloacetate) are endergonic (Fig. 12-3).

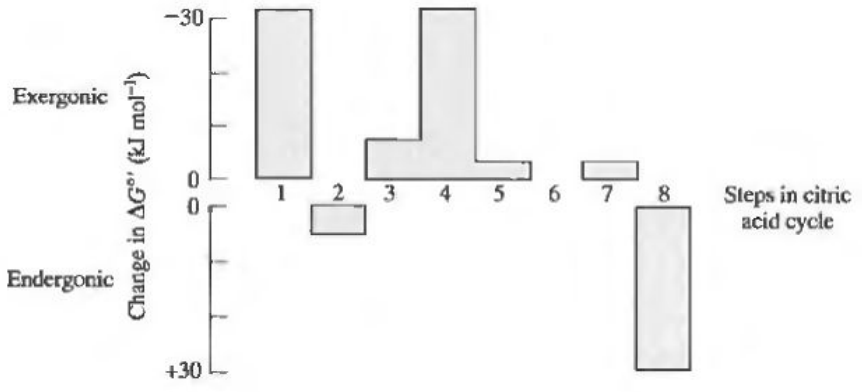


Fig. 12-3 The energetics of the reactions in the citric acid cycle.

Question: How are the two unfavored reactions in the citric acid cycle overcome?

Both of these reactions are followed by exergonic reactions. The equilibrium of the reaction malate $\rightleftharpoons$ oxaloacetate (step 8) lies in favor of malate formation, so at equilibrium the concentration of oxaloacetate will be low. The next reaction in the cycle (oxaloacetate + acetyl-CoA $\rightarrow$ citrate) (step 1) is, however, exergonic, and the oxaloacetate is removed to condense with acetyl-CoA. Similarly, the conversion of citrate to isocitrate is endergonic, and at equilibrium the reaction favors the formation of citrate. The next reaction in the cycle (isocitrate $\rightarrow$ 2-oxoglutarate) is exergonic, and so the isocitrate is removed thus allowing this reaction to proceed.

The rate of utilization of acetyl-CoA in the citric acid cycle depends on the *energy status* within the mitochondria. Under conditions of high bond energy, the concentrations of NADH and ATP are high, and those of NAD^+ and AMP are low. It is usual to describe the energy status, not in terms of absolute concentrations, but in terms of ratios such as NAD^+/NADH and AMP/ATP. So a high-energy status would mean a low NAD^+/NADH ratio or a high NADH/NAD^+ ratio, and a low AMP/ATP ratio or a high ATP/AMP ratio.

EXAMPLE 12.2

What would be the effect on the reactions of the citric acid cycle if the NADH and $FADH_2$ were not reoxidized?

The reactions of the cycle would cease once all the NAD^+ and FAD were reduced to NADH and $FADH_2$, respectively. The reoxidation of NADH and $FADH_2$ occurs in the electron-transport system and is slow in comparison with the *potential* rate at which the citric acid cycle could function. Thus, because the reoxidation of NADH and $FADH_2$ results in energy transduction, the cycle is controlled by the energy requirements of the mitochondria. Remember that the concentrations of these oxidized and reduced cofactors in the cytoplasm and the mitochondria are quite low; they must be rapidly interconverted so that metabolism can proceed.

12.4 REGULATION OF THE CITRIC ACID CYCLE

There are four major regulatory enzymes in the citric acid cycle. These are *citrate synthase* (step 1), *isocitrate dehydrogenase* (step 3), *2-oxoglutarate dehydrogenase* (step 4), and *succinate dehydrogenase* (step 6).

In the first control point, citrate synthase catalyzes the condensation of acetyl-CoA with oxaloacetate to produce citrate ($\Delta G^{\circ\prime} = -32.2\,kJ\,mol^{-1}$). Although the reaction is reversible, the equilibrium lies very much in favor of citrate formation because of the hydrolysis of a bond in the intermediate compound, citroyl-CoA (Fig. 12-4). Citroyl-CoA is bound to citrate synthase, and the hydrolysis of the thioester bond, to produce citrate and coenzyme A, is an exergonic process. Citrate synthase is inhibited by its substrates (acetyl-CoA and oxaloacetate), and its activity is affected by

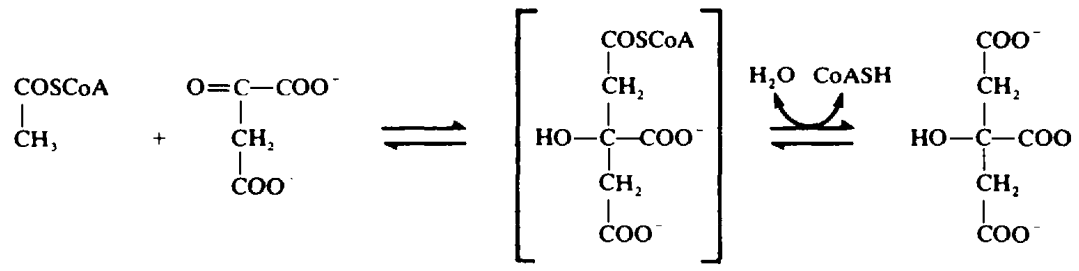


Fig. 12-4 The citroyl-CoA intermediate produced by citrate synthase.

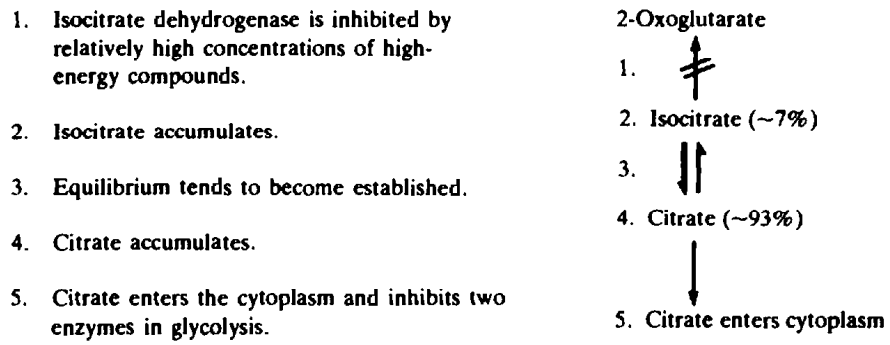


Fig. 12-5 The control of glycolysis by elevated concentrations of high-energy compounds within the mitochondria.

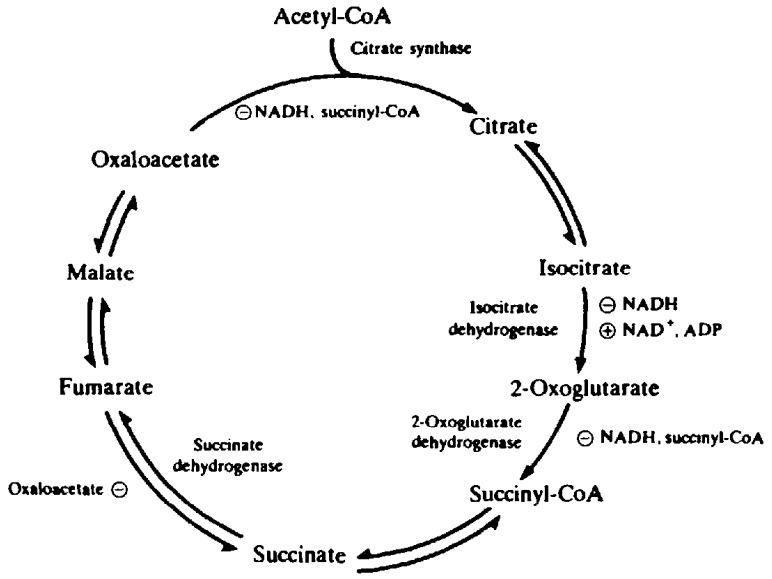


Fig. 12-6 The four regulatory steps in the citric acid cycle. ⊕ = activation; ⊖ = inhibition.

the energy status of the mitochondria (low NAD^+/NADH inhibits) and by succinyl-CoA, which competes with acetyl-CoA for the active site.

The second point of regulation in the cycle is at the conversion of isocitrate to 2-oxoglutarate, catalyzed by isocitrate dehydrogenase. The decarboxylation in this step drives the reaction irreversibly toward the formation of 2-oxoglutarate. Isocitrate dehydrogenase is an allosteric enzyme (Chap. 9) that is activated by ADP and NAD^+ and inhibited by ATP and NADH. When the concentration of compounds with high-energy bonds (or those with the ability to produce compounds with high-energy bonds) rises, the enzyme is inhibited; thus, the remaining reactions of the cycle are impeded, and an accumulation of isocitrate occurs. The production of isocitrate from citrate is a reversible reaction, and if the isocitrate is not removed by other reactions, the citrate isocitrate reaction tends toward equilibrium. At equilibrium, the relative amounts of these two components are ~93 percent and ~7 percent, respectively. So the inhibition of isocitrate dehydrogenase results in an increase in the concentration of citrate, which can leave the mitochondria via a specific transport protein and enter the cytoplasm. This is the source of the citrate that is the inhibitor of the two glycolytic enzymes, phosphofructokinase and pyruvate kinase (Fig. 12-5).

The third control step in the citric acid cycle is catalyzed by 2-oxoglutarate dehydrogenase. This multienzyme complex is subject to *product inhibition* by both NADH and succinyl-CoA. Yet again,

the presence of a high-energy state (high $NADH/NAD^+$) acts to slow the cycle at this reaction.

The fourth control in the cycle is at the conversion of succinate to fumarate via succinate dehydrogenase. This enzyme is inhibited by oxaloacetate, so that if for any reason oxaloacetate accumulates, the enzymes will be inhibited; thus, oxaloacetate *feeds back* and inhibits a reaction that is required for its synthesis. This phenomenon is called *negative feedback*. (See Chap. 9.)

The major regulatory sites in the citric acid cycle are shown in Fig. 12-6.

12.5 THE PYRUVATE DEHYDROGENASE COMPLEX

Acetyl-CoA is produced from fatty acids, proteins, and carbohydrates and is a central and major compound in intermediary metabolism. The mechanism of its formation from the degradation of fatty acids and proteins is discussed in Chaps. 13 and 15, respectively; here, the means whereby *carbohydrates* form this most important molecule will be presented. The glycolytic pathway can yield pyruvate from all degradable sugars, and this can be converted to acetyl-CoA. Pyruvate enters the mitochondrial matrix and is the substrate for the multienzyme complex pyruvate dehydrogenase.

Question: What is the overall reaction catalyzed by the pyruvate dehydrogenase complex?

$$\underset{\text{Pyruvate}}{CH_3COCOO^-} + NAD^+ + CoASH \rightarrow \underset{\text{Acetyl-CoA}}{CH_3COSCoA} + CO_2 + NADH$$

The complex consists of three enzymes: *pyruvate decarboxylase* (E_1), *dihydrolipoyl transacetylase* (E_2), and *dihydrolipoyl dehydrogenase* (E_3). The first reaction is a decarboxylation of pyruvate to produce a 2-hydroxyethyl residue that is bonded to pyruvate dehydrogenase via its prosthetic group *thiamine pyrophosphate* (TPP). The second enzyme of the complex (dihydrolipoyl transacetylase) catalyzes two steps, namely, the transfer of an acetyl group (derived from the oxidation of the 2-hydroxyethyl residue) to a lipoic acid *arm* of the enzyme and then transfer of this acetyl group to CoA to form acetyl-CoA. At this step the lipoic acid is converted to its fully reduced state; it then reduces the FAD prosthetic group of dihydrolipoyl dehydrogenase to $FADH_2$. Finally, the reduced prosthetic group, $FADH_2$, is reoxidized to FAD by converting NAD^+ to NADH. The overall sequence of reactions is operationally irreversible because of the decarboxylation step, which is irreversible by virtue of the relatively low CO_2 concentration in tissues. However, all the *other* reactions are reversible. The various reactions of the complex are given in Fig. 12-7.

Since a major function of the citric acid cycle is to oxidize acetyl-CoA, with the subsequent generation of ATP, the rate at which the cycle of reactions operates depends on the availability of acetyl-CoA. Thus, the *energy status* of the cell exerts control over the activity of the pyruvate dehydrogenase complex. If the concentration of acetyl-CoA or ATP is low, then the complex is activated to produce acetyl-CoA. If the concentration of acetyl-CoA is high, because of either the breakdown of fatty acids or protein, or the existence of a high-energy state within the mitochondria, then the complex is inhibited. This latter control reduces the rate of pyruvate degradation which in turn reduces the rate of glucose degradation.

EXAMPLE 12.3

Control of the activity of the pyruvate dehydrogenase complex is exerted by the phosphorylation of pyruvate decarboxylase (E_1), which renders it *inactive*. This process is catalyzed by *pyruvate dehydrogenase kinase*, which is always tightly bound to E_1. The kinase is activated by high-energy conditions, and it requires ATP to accomplish the phosphorylation step. Another enzyme, *phosphoprotein phosphatase*, is weakly bound to E_1 and reactivates the system by removing the inhibitory phosphate group (Fig. 12-8).

The pyruvate dehydrogenase complex is *not directly* a part of the reactions that constitute the citric acid cycle. It is the *link* between glycolysis and the citric acid cycle, and its activity is controlled by the energy status of the mitochondria.

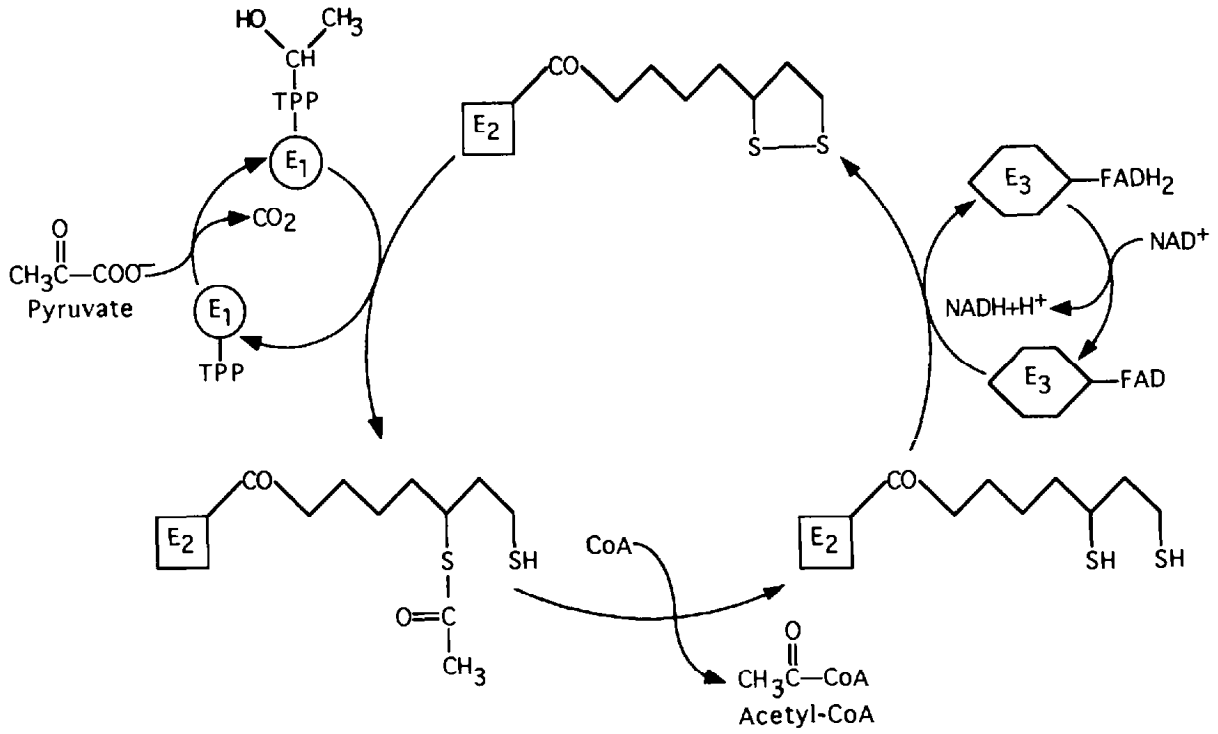
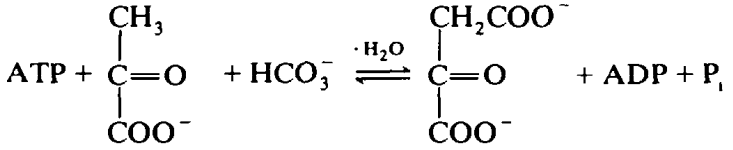


Fig. 12-7 The reactions of the pyruvate dehydrogenase complex. The reactants in the overall reaction are shown in boxes. E_1 = pyruvate decarboxylase (TPP = thiamine pyrophosphate as prosthetic group), E_2 = dihydrolipoyl transacetylase (oxidized lipoic acid as prosthetic group), E_3 = dihydrolipoyl dehydrogenase (FAD as prosthetic group).

12.6 PYRUVATE CARBOXYLASE

Pyruvate carboxylase is another enzyme which is *not* a part of the citric acid cycle per se but which functions in close association with it. The function of this enzyme is described in Chap. 11, but it is useful to consider its action, and that of pyruvate dehydrogenase, in relation to the citric acid cycle.



Pyruvate **Oxaloacetate**

The product of this reaction, oxaloacetate, can either enter the gluconeogenic pathway (Chap. 11) by way of malate or condense with acetyl-CoA to yield citrate. Pyruvate carboxylase is an allosteric enzyme, and it is activated by the heterotropic effector, acetyl-CoA. Thus, pyruvate in the mitochondria is the substrate for either pyruvate dehydrogenase or pyruvate carboxylase, the activities of which, in turn, are controlled by reactants associated with the citric acid cycle. The interplay among pyruvate dehydrogenase, pyruvate carboxylase, pyruvate, and the citric acid cycle is shown in Fig. 12-9.

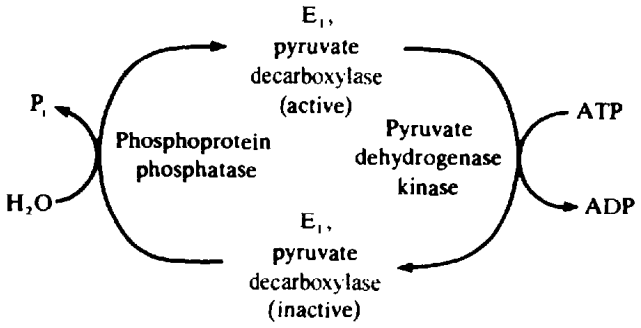


Fig. 12-8 The control of the pyruvate dehydrogenase complex via phosphorylation/dephosphorylation of the pyruvate decarboxylase part of the complex.

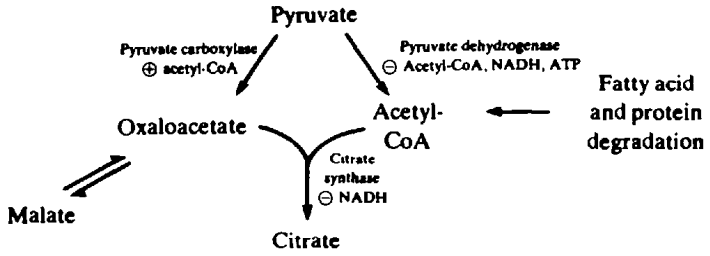


Fig. 12-9 The fates of pyruvate in the mitochondria. ⊕ = activation; ⊖ = inhibition.

EXAMPLE 12.4

The fate of pyruvate within the mitochondria depends on the energy status that exists in the mitochondria. If the levels of high-energy compounds are elevated, i.e., ATP and NADH are plentiful, then pyruvate dehydrogenase will be inhibited. If the concentration of acetyl-CoA is high (perhaps from the degradation of fatty acids or protein), then this will also tend to inhibit pyruvate dehydrogenase and pyruvate will be conserved through inhibition of its conversion to acetyl-CoA. But under these conditions, the acetyl-CoA causes the activation of pyruvate carboxylase, and so oxaloacetate is produced. A lack of oxaloacetate precludes the conversion of acetyl-CoA to citrate, resulting in a high level of acetyl-CoA. By activating pyruvate carboxylase, this mechanism permits the necessary production of oxaloacetate from pyruvate. (Other reasons for a lack of oxaloacetate are discussed in Example 12.6.)

12.7 THE AMPHIBOLIC NATURE OF THE CITRIC ACID CYCLE

Apart from the production of NADH and $FADH_2$, which are the high-energy fuels of electron transport, the citric acid cycle has two other major functions. Several of its intermediate compounds are used to synthesize other cell constituents. This, the provision of molecules for other metabolic or biosynthetic pathways, is the *anabolic* function of the cycle (Table 12.1). Alternatively, certain other processes occurring within the cell may produce intermediates of the citric acid cycle. These compounds enter the reactions of the cycle, and their degradation involves the *catabolic* role of the cycle. These two major capabilities classify the citric acid cycle as an *amphibolic* pathway (Greek: *amphi* meaning "both sides").

Table 12.1. Anabolic and Catabolic Reactions of the Citric Acid Cycle

Anabolic Reactions	
Intermediate Removed	Metabolic Fate
Citrate	Fatty acid biosynthesis
2-Oxoglutarate	Synthesis of glutamate
Succinyl-CoA	Heme biosynthesis
Malate	Gluconeogenesis
Oxaloacetate	Synthesis of aspartate
Catabolic Reactions	
Intermediate Produced	Source
2-Oxoglutarate	Glutamate
Succinyl-CoA, Oxaloacetate, Fumarate, 2-Oxoglutarate	Degradation of some amino acids (Chap. 15)
Oxaloacetate	Aspartate

EXAMPLE 12.5

If intermediates of the citric acid cycle enter the cycle via other reactions in the cell, are they oxidized?

If one molecule of 2-oxoglutarate, for example, were to enter the cycle, its metabolism through one turn of the cycle would be:

$$\text{2-Oxoglutarate}^{2-} + 3\text{NAD}^+ + \text{FAD} + \text{GDP}^{3-} + \text{P}_i^{2-} + 2\text{H}_2\text{O} + \text{Acetyl-CoA} \rightarrow \text{2-Oxoglutarate}^{2-} + 2\text{CO}_2 + 3\text{NADH} + \text{FADH}_2 + 2\text{H}^+ + \text{GTP}^{4-} + \text{CoA}$$

Only the acetyl group of acetyl-CoA is oxidized to CO_2; the 2-oxoglutarate is regenerated.

EXAMPLE 12.6

If intermediates pass from the cycle and enter other anabolic pathways, why does the cycle continue to operate?

The cycle oxidizes acetyl-CoA, and to perform this task, it must convert acetyl-CoA to citrate. For this to be achieved, oxaloacetate must be available. If the removal of intermediates results in a decrease in the amount of oxaloacetate for this purpose, acetyl-CoA cannot be removed and will accumulate. This will *inhibit* the pyruvate dehydrogenase complex and *activate* pyruvate carboxylase, leading to the conversion of pyruvate to oxaloacetate. This product is now available to condense with the acetyl-CoA to produce citrate, which will restore the status quo. Reactions like that of pyruvate carboxylase that provide molecules for the replacement of intermediates of the citric acid cycle are known as *anaplerotic* reactions (Greek, meaning "to fill up": *ana* = "up" + *plerotikos* from *pleroun* = "to make full").

12.8 THE GLYOXYLATE CYCLE

Some plants and bacteria that can use acetate as their sole source of carbon are able to oxidize acetyl-CoA via the citric acid cycle, or the acetate can be converted to carbohydrates via a pathway that is a modification of the citric acid cycle. This pathway is known as the *glyoxylate cycle* (Fig. 12-10)

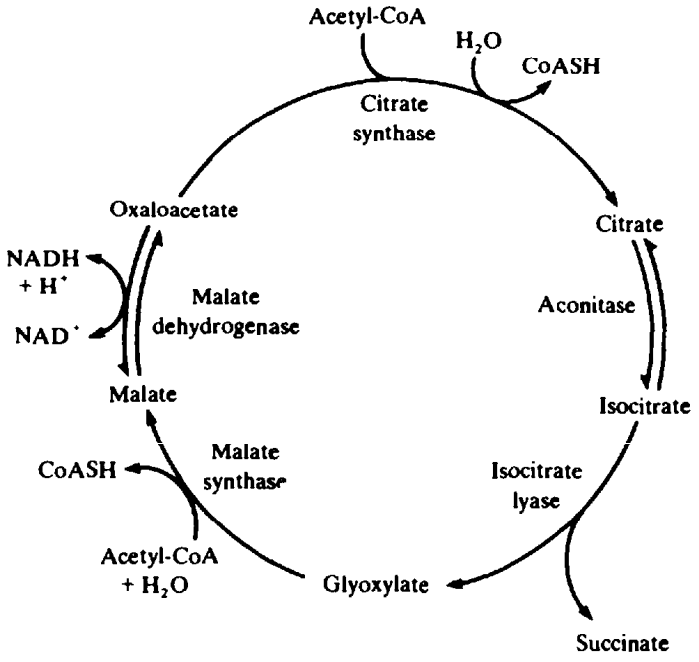
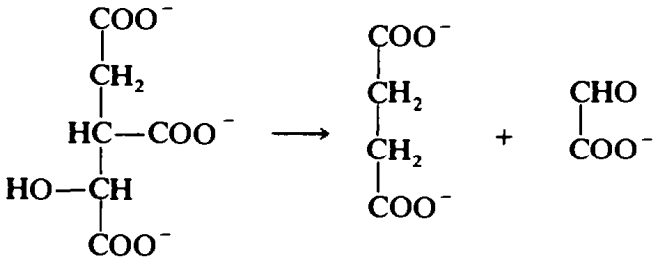
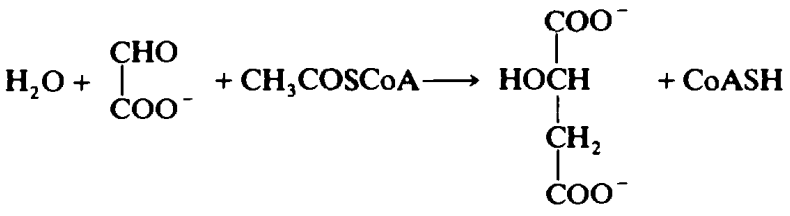


Fig. 12-10 The glyoxylate cycle.

and in plants it exists in organelles called *glyoxysomes*. The enzymes of the cycle are very active in germinating plant seeds, which are able to convert their stores of fatty acids into carbohydrates during their growth. Animals do not have the enzymes of the glyoxylate cycle and therefore cannot convert fatty acids into carbohydrates. The two reactions unique to this pathway are catalyzed by the enzymes *isocitrate lyase* and *malate synthase*.



Isocitrate Succinate Glyoxylate



Glyoxylate Acetyl-CoA Malate Coenzyme A

These reactions produce two important intermediate compounds, *succinate* and *malate* (which is converted into *oxaloacetate*). The two decarboxylation steps of the citric acid cycle are bypassed, and so there is no oxidation of acetyl-CoA to CO_2. Two molecules of acetyl-CoA are used, but *all* the carbon atoms are retained.

Question: What is the overall reaction of the glyoxylate cycle?

$$2\ \text{Acetyl-CoA} + 2H_2O + NAD^+ \rightarrow \text{Succinate}^{2-} + NADH + 3H^+ + 2\ \text{CoASH}$$

Solved Problems

REACTIONS OF THE CITRIC ACID CYCLE

12.1. What are the overall chemical changes that occur during one complete turn of the citric acid cycle?

SOLUTION

The overall reactions are the complete oxidation of one molecule of acetyl-CoA, the production of two molecules of CO_2, the reduction of three molecules of NAD^+ and one of FAD, and the phosphorylation of one molecule of GDP.

12.2. The overall reaction of the citric acid cycle is

$$CH_3COSCoA + 2H_2O + 3NAD^+ + FAD + GDP + P_i \rightarrow 2CO_2 + 3NADH + 3H^+ + FADH_2 + GTP$$

Is only one ATP equivalent (GTP) produced for each molecule of acetyl-CoA consumed?

SOLUTION

Only 1 ATP equivalent is *directly* produced, but the reoxidation of NADH and $FADH_2$ yields a further 11 molecules of ATP by electron transport and oxidative phosphorylation (Chap. 14).

12.3. In the citric acid cycle, how many steps involve (*a*) oxidation-reduction, (*b*) hydration-dehydration, (*c*) substrate-level phosphorylation, and (*d*) decarboxylation? List the enzymes responsible for these reactions.

SOLUTION

(*a*) Four steps involve oxidation-reduction. The enzymes involved are isocitrate dehydrogenase, 2-oxoglutarate dehydrogenase, succinate dehydrogenase, and malate dehydrogenase.

(*b*) Two steps involve hydration-dehydration reactions. The enzymes responsible for these reactions are aconitase and fumarase.

(*c*) One step involves substrate-level phosphorylation, and the enzyme is succinyl-CoA synthetase.

(*d*) Two steps involve decarboxylation. The enzymes are isocitrate dehydrogenase and 2-oxoglutarate dehydrogenase.

12.4. Two molecules of CO_2 are produced each time a molecule of acetyl-CoA is oxidized. Do the carbon atoms of *this* acetyl-CoA molecule become the CO_2 in the first turn of the cycle?

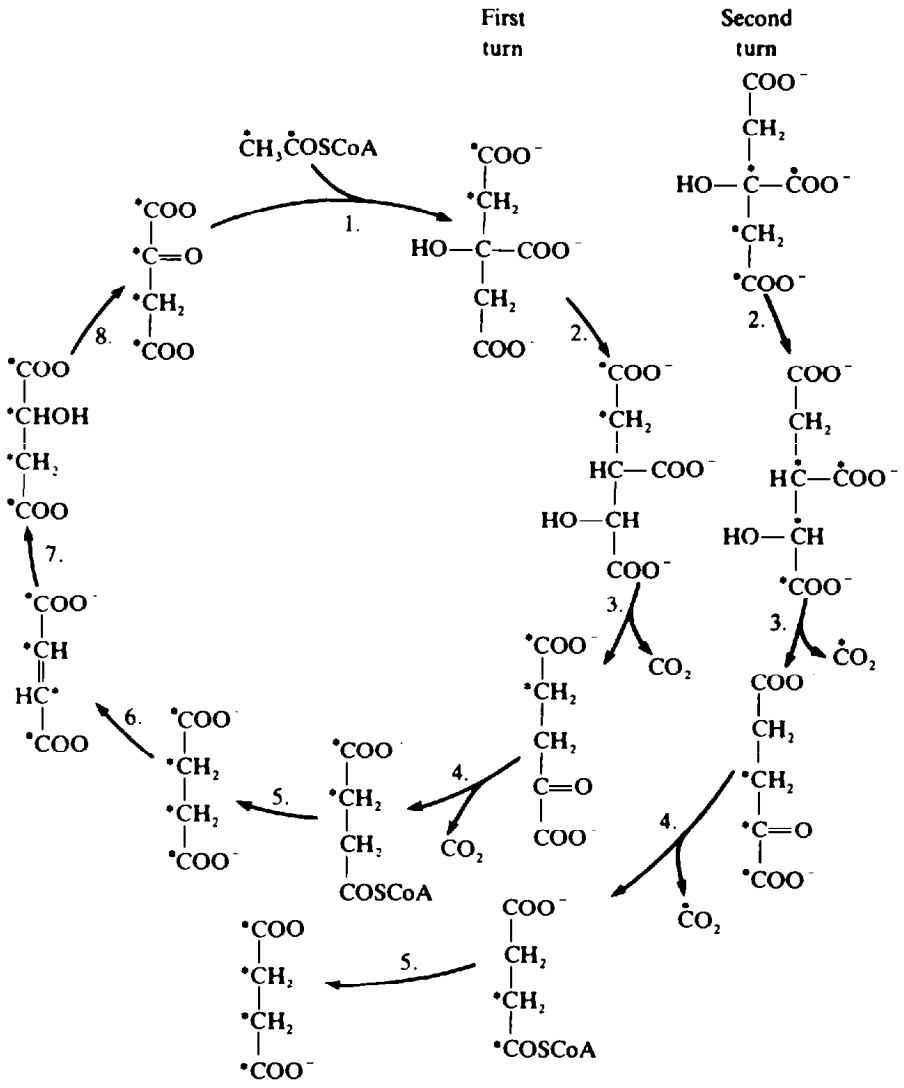


Fig. 12-11 The distribution of the carbon atoms of acetyl-CoA in the citric acid cycle.

SOLUTION

No. Figure 12-11 shows the fate of the carbon atoms of one molecule of acetyl-CoA in *two* turns of the cycle. The two carbon atoms can be followed until step 5, the formation of the *symmetrical* molecule succinate, when they become *randomized*. This means that the two methylene carbons of the symmetrical molecule succinate have equal probability of arising from the methyl group of acetyl-CoA. Steps 3 and 4 of the *second* turn each yield CO_2, which represents, at each step, 50 percent of the original carboxyl carbon of acetyl-CoA. During the *third* turn of the cycle, these two steps collectively liberate 50 percent of the remaining labeled methyl carbon as CO_2, and so forth for subsequent turns of the cycle.

12.5. Citrate is a symmetrical molecule, yet in step 2 of the citric acid cycle, aconitase catalyzes the removal of the elements of water from only *half* of the molecule and not from the identical other half. How is this rationalized?

SOLUTION

Hans Krebs, the discoverer of the citric acid cycle, also pondered this question and at one stage came to the conclusion that citrate was not an intermediate in the cycle. However, in 1948 Alexander

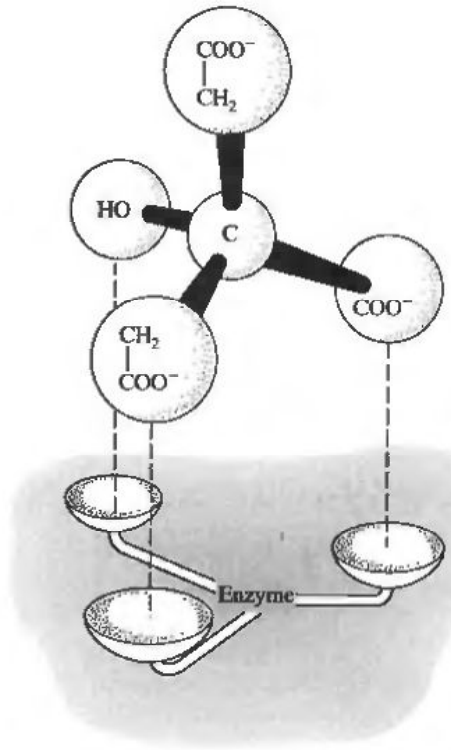


Fig. 12-12 The three-point attachment of citrate to aconitase. This shows how only one *particular* $-CH_2COO^-$ group binds to the enzyme.

Ogston offered an explanation, called the *three-point attachment* proposal, that was to initiate the concept of *prochirality*. If citrate is represented as a three-dimensional structure (Fig. 12-12), then on the assumption that a three-point attachment to aconitase is necessary for catalysis, it is apparent that citrate can only be accommodated in one orientation. The removal of the elements of water can then only occur from one particular half of the symmetrical molecule.

12.6. If a cell in which the citric acid cycle operates has the enzyme *phosphoenolpyruvate carboxykinase*, by what processes might it oxidize one molecule of 2-oxoglutarate to five molecules of CO_2?

SOLUTION

The theoretical reaction sequence would be

$$\text{2-Oxoglutarate} \rightarrow (CO_2) + \text{Succinyl-CoA} \rightarrow \text{Fumarate} \rightarrow \text{Malate} \rightarrow \text{Oxaloacetate} \rightarrow (CO_2) + \text{Phosphoenolpyruvate} \rightarrow \text{Pyruvate} \rightarrow (CO_2) + \text{Acetyl-CoA} \rightarrow 2CO_2$$

12.7. 2-Fluoroacetate is an animal poison found in some South African plants; it is also used in baits to kill wild rabbits (the poison called "1080"). When ingested, it is converted into 2-fluoroacetyl-CoA. What is the fate of 2-fluoroacetyl-CoA in the citric acid cycle and what is the basis of the toxicity of 2-fluoroacetate?

SOLUTION

2-Fluoroacetyl-CoA competes with acetyl-CoA as a substrate for citrate synthase and is converted into 4-fluorocitrate. It is 4-fluorocitrate that blocks the citric acid cycle by being a potent inhibitor of aconitase (Fig. 12-13) and in this lies its toxicity.

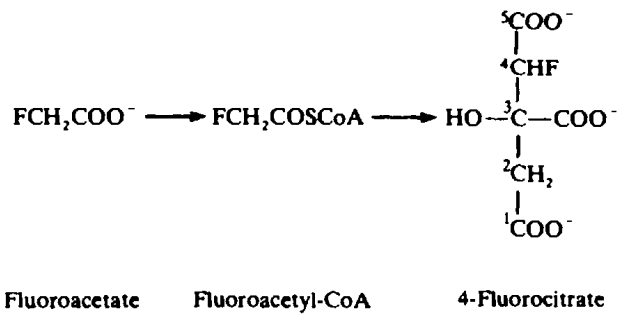


Fig. 12-13 The synthesis of 4-fluorocitrate from 2-fluoroacetate.

THE PYRUVATE DEHYDROGENASE COMPLEX AND PYRUVATE CARBOXYLASE

12.8. By which reactions is pyruvate converted to succinate without depleting any of the intermediates of the citric acid cycle?

SOLUTION

Pyruvate can be converted to acetyl-CoA via the pyruvate dehydrogenase complex. Pyruvate can also be carboxylated via pyruvate carboxylase to produce oxaloacetate. So, two molecules of pyruvate can form the precursors of citrate, which can be converted to succinate within the citric acid cycle.

THE AMPHIBOLIC NATURE OF THE CITRIC ACID CYCLE

12.9. Assuming aspartate is the only major carbon source supplied to a preparation of cells, outline all reactions whereby the citric acid cycle would operate in the mitochondria.

SOLUTION

Aspartate would be *transaminated* (Chap. 15) to yield oxaloacetate. Oxaloacetate would then be converted to acetyl-CoA by the following sequence of reactions:

$$\text{Oxaloacetate} \xrightarrow{1} \text{Malate} \xrightarrow{2} \text{Oxaloacetate} \xrightarrow{3} \text{Phosphoenolpyruvate} \xrightarrow{4} \text{Acetyl-CoA}$$

The enzymes are, respectively: (1) mitochondrial malate dehydrogenase; (2) cytoplasmic malate dehydrogenase; (3) phosphoenolpyruvate carboxykinase; (4) pyruvate kinase and pyruvate dehydrogenase. The acetyl-CoA could then condense with oxaloacetate (produced from a second molecule of aspartate) to yield citrate. Aspartate could, therefore, continue to supply acetyl-CoA, which would continue to fuel the citric acid cycle.

THE GLYOXYLATE CYCLE

12.10. The glyoxylate and citric acid cycles have several reactions in common, but what two enzymes are unique to the glyoxylate cycle?

SOLUTION

Isocitrate lyase and malate synthase. Isocitrate lyase catalyzes the cleavage of isocitrate to succinate and glyoxylate; malate synthase catalyzes the condensation of glyoxylate and acetyl-CoA to yield malate.

12.11. Animals are unable to achieve a net gain of carbohydrate from acetyl-CoA, whereas plants can. What is the explanation for this?

SOLUTION

The formation of acetyl-CoA from pyruvate in animals is via the pyruvate dehydrogenase complex, which catalyzes the irreversible decarboxylation reaction. Carbohydrate is synthesized from oxaloacetate, which in turn is synthesized from pyruvate via pyruvate carboxylase. Since the pyruvate dehydrogenase reaction is irreversible, acetyl-CoA cannot be converted to pyruvate, and hence animals cannot realize a net gain of carbohydrate from acetyl-CoA. Because plants have a glyoxylate cycle and animals do not, plants synthesize one molecule of succinate and one molecule of malate from two molecules of acetyl-CoA and one of oxaloacetate. The malate is converted to oxaloacetate, which reacts with another molecule of acetyl-CoA and thereby continues the reactions of the glyoxylate cycle. The succinate is also converted to oxaloacetate via the enzymes of the citric acid cycle. Thus, one molecule of oxaloacetate is *diverted* to carbohydrate synthesis and, therefore, plants are able to achieve net synthesis of carbohydrate from acetyl-CoA.

Supplementary Problems

12.12. Calculate the number of molecules of ATP that can be produced by substrate-level phosphorylation of ADP from two molecules of acetyl-CoA in the (*a*) citric acid cycle and (*b*) glyoxylate cycle.

12.13. Fumarate contains an *E* (or *trans*) double bond. What would be the result if a cell were supplied with the *Z* (or *cis*) isomer of fumarate (maleate) as the sole carbon source?

12.14. What effect, if any, would a lack of oxygen have on the rate of the citric acid cycle in red blood cells?

12.15. If a potent inhibitor of succinyl-CoA synthetase is applied to liver cells and germinating plant cells, what effect would this have on energy production and the synthesis of carbohydrate in both types of cell?

Chapter 13

Lipid Metabolism

13.1 INTRODUCTION

The major dietary lipids for humans are animal and plant *triacylglycerols*, *sterols*, and *membrane phospholipids*. The process of lipid metabolism fashions and degrades the lipid stores and produces the structural and functional lipids characteristic of individual tissues. For example, the evolution of a highly organized nervous system has depended on the natural selection of specific enzymes to synthesize and degrade (turn over) the lipids of the brain and central nervous system.

13.2 LIPID DIGESTION

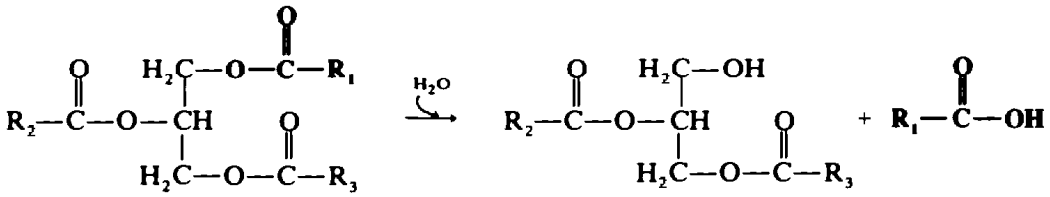
People of western culture ingest about 100 g of triacylglycerol per day. The digestion and absorption of this lipid, together with the ingested phospholipids, depend on secretions from the pancreas (exocrine) and a flow of bile from the gall-bladder. The important constituents of the pancreatic secretions are enzymes, and those of the bile are the bile salts (Chap. 6).

The Enzymes

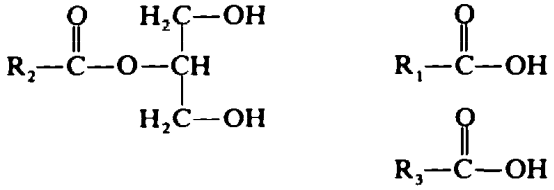
Lipid digestion is accomplished in the small intestine by the action of hydrolytic enzymes, called *lipases* and *phospholipases*, which act on dietary triacylglycerol and phospholipids, respectively.

EXAMPLE 13.1

Consider the action of a lipase:



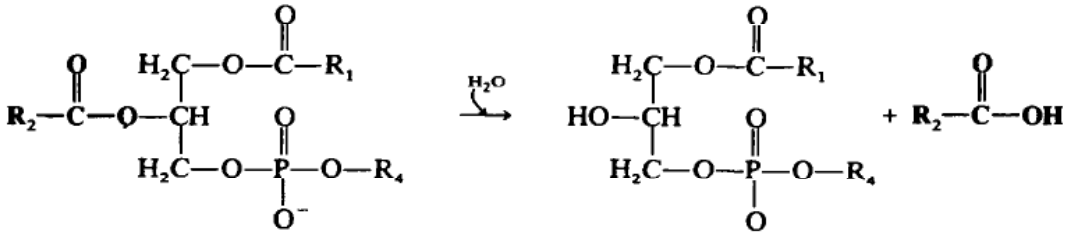
R_1, R_2, and R_3 are the hydrocarbon chains of fatty acids. Ester bonds between a fatty acid and glycerol are hydrolyzed. In the lumen of the intestine, the complete action of pancreatic lipase on dietary triacylglycerol produces the following:



2-Monoacylglycerol and 2 Fatty acids

EXAMPLE 13.2

Phospholipase A_2 catalyzes the following reaction:



R_1 and R_2 are the carbon chains of fatty acids, R_4 is an alcohol. Only one ester bond between fatty acid and glycerol is hydrolyzed, specifically at position 2 of the glycerol carbon chain.

Phospholipase A_1 hydrolyzes the ester bond between a fatty acid and glycerol in position 1 of the carbon chain of a phosphoglyceride.

Question: How do hydrolytic enzymes act on lipids that repel water?

It is necessary for these enzymes to act at a water-lipid interface. Digestive lipases secreted into the lumen of the small intestine associate with the surface of large fat droplets. The initial products of digestion, by lipases and phospholipases, are fatty acids and lysophosphoglycerides, which are strong *detergents*. These hasten the digestive process because they disperse the large fat droplets into myriads of tiny ones. As the concentration of fatty acids increases and 2-monoacylglycerol is produced, these are incorporated into micelles (Chap. 6) of bile salts. Monoacylglycerol also increases the detergent action of the bile salts, thus also facilitating emulsification of triacylglycerol and lipid-soluble vitamins. The mixed micelles migrate in large numbers to the surface of the intestinal epithelial cells, where the fatty acids, lipid-soluble vitamins and 2-monoacylglycerol are released from the micelle (Fig. 13-1).

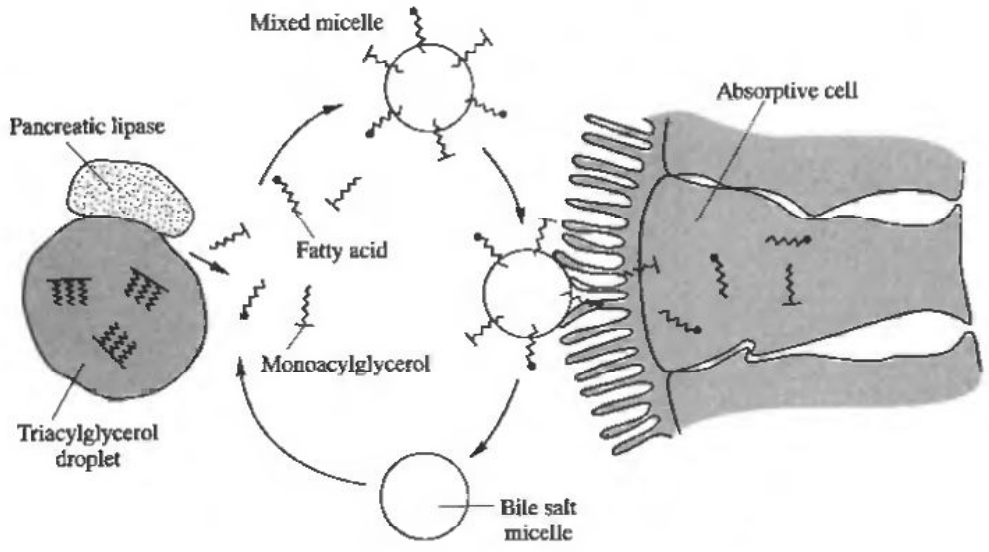


Fig. 13-1 Digestion of triacylglycerol in the intestine.

Absorption of Fatty Acids

Fatty acids of carbon chain length equal to, or greater than, 14 (long-chain acids) diffuse passively into the intestinal epithelial cells. They enter the cells down a concentration gradient because the

concentration of free fatty acids in the mixed micelle is high and that of the cell is low. The cell membrane is no barrier to the lipophilic fatty acid. Entry of fatty acid into the cell is immediately followed by binding to a *binding protein*, which has a high affinity for long-chain fatty acids. Simultaneously the 2-monoacylglycerol passively diffuses into the epithelial cell and, with the fatty acids, is converted rapidly to triacylglycerol.

Question: By what mechanism do the fatty acids and 2-monoacylglycerol enter the circulation?

The newly synthesized triacylglycerol becomes organized into *chylomicrons* (a type of lipoprotein; see next section), which are secreted by the intestinal epithelial cell into the *lacteals*, small lymph vessels in the villi of the small intestine. Then from the *lymphatics*, the chylomicrons pass into the *thoracic duct*, from which they enter the blood and thus contribute to the transport of lipid fuel to various tissues. A feature of chylomicron metabolism is their ability to deliver lipid fuels to extrahepatic tissues.

13.3 LIPOPROTEIN METABOLISM

Role of Lipoproteins

Lipoproteins transport hydrophobic fats in plasma (Fig. 13-2). The major lipoproteins (Chap. 6) circulating in the blood are chylomicrons, VLDLs (very low density lipoproteins), LDLs (low-density lipoproteins), and HDLs (high-density lipoproteins). IDLs (intermediate density lipoproteins) are derived from VLDLs in the formation of LDLs. Fatty acids are important cellular

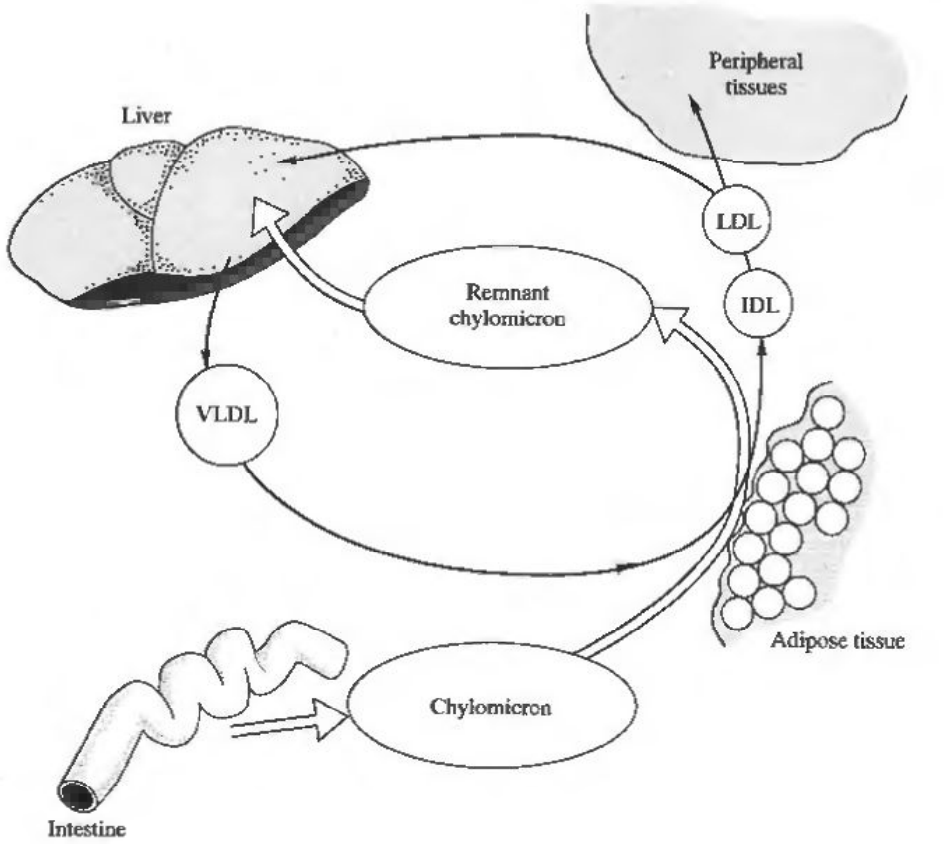


Fig. 13-2 Circulation of lipoproteins in the blood and tissues.

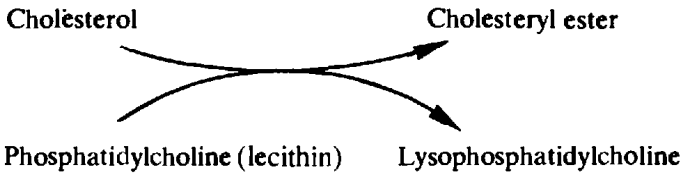
fuels and are stored as triacylglycerols in *adipose tissue*. Fatty acids destined for storage as depot fat are transported to adipose tissue principally as triacylglycerol in chylomicrons and VLDLs. In adipose tissue, chylomicrons are rapidly degraded, and the *remnant* particles reenter the circulation and are taken up by the liver. VLDLs are degraded in adipose tissue to LDLs which then circulate as the major transport lipoproteins for cholesterol. HDLs are lipoproteins that continuously circulate; they contain an enzyme, *phosphatidylcholine:cholesterol acyltransferase* (or *lecithin:cholesterol acyltransferase*; LCAT), that converts free cholesterol to cholesteryl esters.

Question: What is the fate of the cholesteryl esters in HDL?

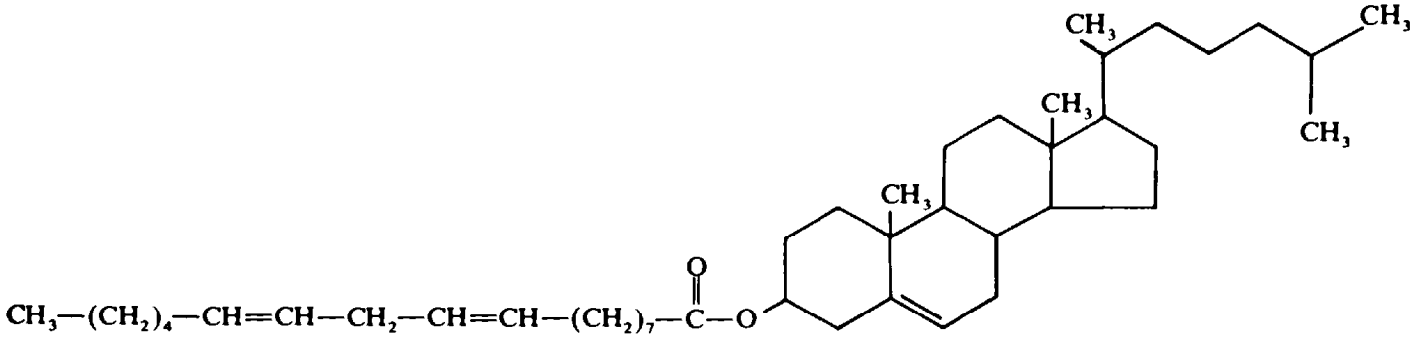
Cholesteryl esters are transferred out of HDL by cholesteryl ester transfer protein (CETP). CETP promotes the transfer of cholesteryl esters to VLDL and LDL in exchange for triacylglycerol. In this way, CETP enables HDL to transport more cholesteryl esters derived from the LCAT reaction.

EXAMPLE 13.3

The reaction catalyzed by LCAT is:



Linoleic acid is the fatty acid most commonly transferred from phosphatidylcholine to cholesterol, forming the cholesteryl ester, linoleoylcholesterol:



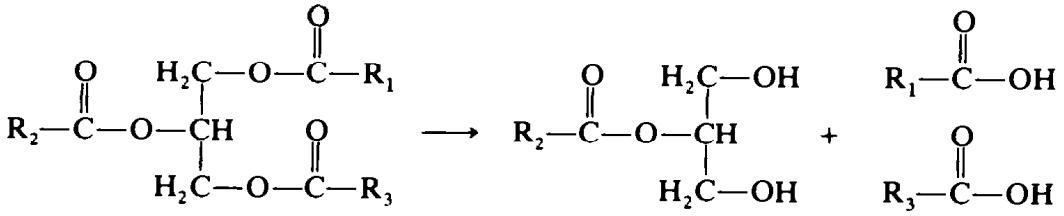
Linoleoylcholesterol

Release of Fatty Acids from Lipoproteins

The action of an enzyme *lipoprotein lipase* in adipose tissue depletes chylomicrons and VLDLs of their triacylglycerol. The enzyme is activated by *apolipoprotein C*, with which it specifically interacts on the surface of chylomicrons and VLDLs.

Question: What is the reaction catalyzed by lipoprotein lipase?

Lipoprotein lipase is an *extracellular* enzyme that hydrolyzes triacylglycerol into 2-monoacylglycerol and two fatty acids. The fatty acids then enter the cell passively down a concentration gradient.



In adipose tissue, triacylglycerol is resynthesized from the fatty acids and stored in a large fat droplet, occupying up to 96 percent of the cellular space in a fat cell (Chap. 1). The fat stores in adipose tissue of an average 70-kg person are sufficient to satisfy the body's energy needs over a long period of starvation (~40 days). Lipoprotein lipase also occurs as an extracellular enzyme in other tissues, and the fatty acids, obtained from the hydrolysis of lipoprotein triglyceride, serve as an immediate cellular fuel, or may be stored as triglyceride.

Release of Cholesterol from LDL

When LDL is abundant in the circulation it provides tissues with an exogenous source of cholesterol.

Question: How is the cholesterol in LDL transferred into cells?

LDL binds specifically to *lipoprotein receptors* on the cell surface. The resulting complexes become clustered in regions of the plasma membrane called *coated pits*. *Endocytosis* follows (see Fig. 13-3). The *clathrin coat* dissociates from the endocytic vesicles, which may recycle the receptors to the plasma membrane or fuse with *lysosomes*. The lysosomal proteases and lipases then catalyze the hydrolysis of the LDL-receptor complexes; the protein is degraded completely to amino acids, and cholesteryl esters are hydrolyzed to free cholesterol and fatty acid. New LDL receptors are synthesized on the endoplasmic reticulum (ER) membrane and are subsequently reintroduced into the plasma membrane. The cholesterol is incorporated in small amounts into the endoplasmic reticulum membrane or may be stored after esterification as cholesteryl ester in the cytosol; this occurs if the supply of cholesterol exceeds its utilization in membranes. Normally, only very small amounts of cholesteryl ester reside inside cells, and the majority of the free cholesterol is in the plasma membrane.

EXAMPLE 13.4

Acyl-CoA:cholesterol acyltransferase catalyzes the esterification of cholesterol in the cell cytoplasm:

The LDL Receptor

The LDL receptor protein has an apparent molecular weight of 160,000. It is an integral membrane glycoprotein consisting of 839 residues folded into five domains. Domain 1 is the ligand binding domain which mediates the interaction with apolipoprotein B or apolipoprotein E. It is rich in cysteines and is negatively charged. Domain 2 has a high degree of homology with the precursor of epidermal growth factor (EGF). The function of this domain is not known, but it may have a

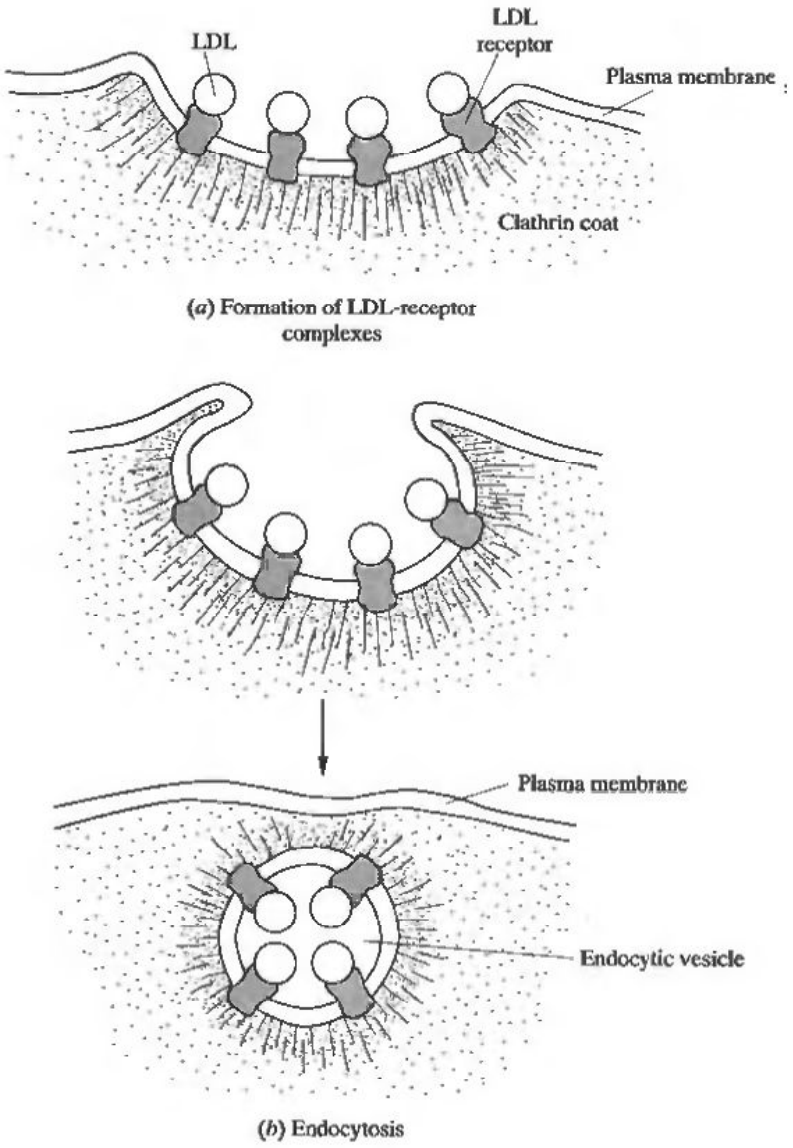


Fig. 13-3 The fate of LDL-receptor complexes in cholesterol uptake into cells.

supportive role in the binding of LDL. Domain 3 contains *O*-linked carbohydrate chains and although its function is not clear, it may act as a "stalk" which separates the binding site from the cell membrane. Domain 4 is a *transmembrane domain* which anchors the receptor into the membrane. Domain 5 is a short cytoplasmic region which targets LDL receptors to coated pits.

Four classes of LDL receptor mutations have been identified. Class 1 mutations are characterized by the failure of expression of the receptor protein. It is possible, however, that a modified protein is produced but it is not recognized as an LDL receptor protein. Class 2 mutations involve a *nonsense mutation* (premature termination of protein synthesis; Chap. 17), and result in a defect in the transfer of the receptor from the endoplasmic reticulum to the cell membranes. This class of mutation is common in Afrikaners and Lebanese. The *Watanabe heritable hyperlipidemic rabbit* (WHHL) is an animal model which has a Class 2 defect and has been used extensively for the study of *familial hypercholesterolemia*. Class 3 mutations result in abnormal binding of LDL. This can be caused by alterations in the amino acid sequence of Domain 1. Class 4 mutations are those with defective internalization due to the receptor's inability to be located in coated pits. This is the result of mutations in the fifth, C-terminal domain.

Question: What chemical mechanism prevents the excessive accumulation of intracellular cholesterol?

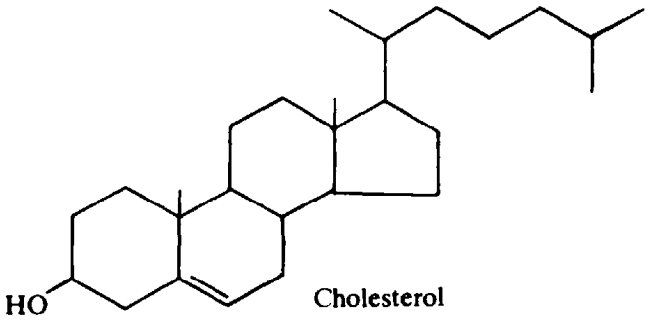
The synthesis of *LDL receptors* is inhibited by an excess of intracellular cholesterol, thus preventing the appearance of new receptors on the cell surface. This may lead to high circulating levels of cholesterol.

EXAMPLE 13.5

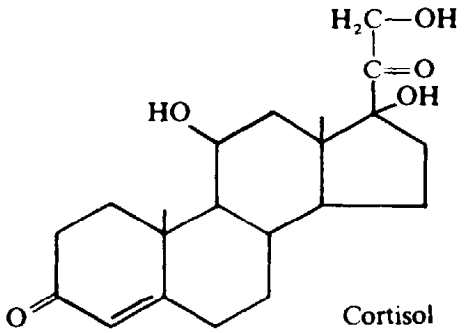
In the congenital disease *familial hypercholesterolemia*, the high circulating level of cholesterol is due to the complete absence of LDL receptors or to the presence of defective receptors on cell surfaces.

EXAMPLE 13.6

Tissues that have a large requirement for cholesterol, such as the adrenal cortex, have a large number of LDL receptors on their cell surfaces. In the case of the adrenal gland, the cholesterol is used in the synthesis of steroid hormones. One such hormone is *cortisol* (hydrocortisone); the similarity of its structure to cholesterol is evident from the following:



Cholesterol



Cortisol

13.4 MOBILIZATION OF DEPOT LIPID

Fatty acids stored as triacylglycerol in adipose tissue are a major source of energy for a variety of tissues when the availability of glucose is low. Stress, prolonged exercise, and starvation lead to the mobilization of depot lipid. Triacylglycerol is hydrolyzed by a so-called *hormone-sensitive lipase*, and fatty acids are released into the circulation. These unesterified fatty acids, bound to *serum albumin*, pass via the circulation to other tissues. By binding to them, the albumin facilitates their entry into the circulation and also minimizes their detergent action during transport.

Question: To which hormones is hormone-sensitive lipase sensitive?

The hydrolysis of triacylglycerol to monoacylglycerol and fatty acids by hormone-sensitive lipase can be stimulated by *epinephrine*, *norepinephrine*, *adrenal steroids*, *glucagon*, and the *hypophysial hormones*, luteotropin (prolactin or luteinizing hormone), β- and α-lipotropins, somatotropin, thyrotropin, and vasopressin.

Question: What is the fate of monoacylglycerol?

A second intracellular lipase, not sensitive to hormones, completes the hydrolysis of monoacylglycerol to glycerol and a fatty acid.

13.5 OXIDATION OF FATTY ACIDS

Oxidation of fatty acids occurs in three well-defined steps; namely, activation, transport into mitochondria, and oxidation to acetyl-CoA.

Fatty Acid Activation

In general, the entry of a fatty acid into a metabolic pathway is preceded by its conversion to its coenzyme A (CoASH) derivative; this acyl derivative is called an *alkanoyl-* or *alkenoyl-CoA*, and in this form the fatty acid is said to have been *activated*.

$$\underset{\text{Hexadecanoate (palmitate)}}{CH_3\text{-}(CH_2)_{14}\text{-}COO^-} \rightarrow \underset{\text{Hexadecanoyl-CoA (palmitoyl-CoA)}}{CH_3\text{-}(CH_2)_{14}\text{-}COSCoA}$$

$$\underset{\text{Octadecenoate (oleate)}}{CH_3\text{-}(CH_2)_7CH = CH\text{-}(CH_2)_7\text{-}COO^-} \rightarrow \underset{\text{Octadecenoyl-CoA (oleoyl-CoA)}}{CH_3\text{-}(CH_2)_7CH = CH\text{-}(CH_2)_7\text{-}COSCoA}$$

The activation of a fatty acid induces the formation of a thioester of fatty acid and CoA. The process is coupled to the hydrolysis of ATP to AMP. For palmitic acid, the reaction is:

$$CH_3\text{-}(CH_2)_{14}\text{-}COO^- + CoASH + ATP \rightarrow CH_3\text{-}(CH_2)_{14}\text{-}COSCoA + AMP + PP_i$$

The enzyme that catalyzes the reaction is *acyl-CoA synthetase*.

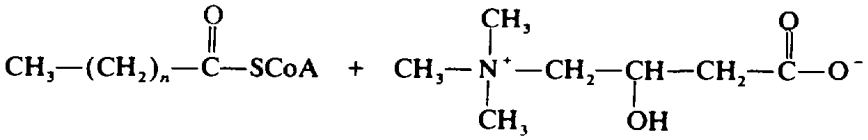
Fatty acids of widely differing chain length can be activated, there being three acyl-CoA synthetase enzymes. One activates acetate (C_2), propionate (C_3), and butyrate (C_4); a second activates medium-chain-length fatty acids (C_4–C_{12}); a third activates long- and medium-chain-length fatty acids. The long-chain acyl-CoA synthetase occurs in the mitochondria and endoplasmic reticulum and is widespread in mammalian tissues.

Transport of Activated Fatty Acids into the Mitochondria

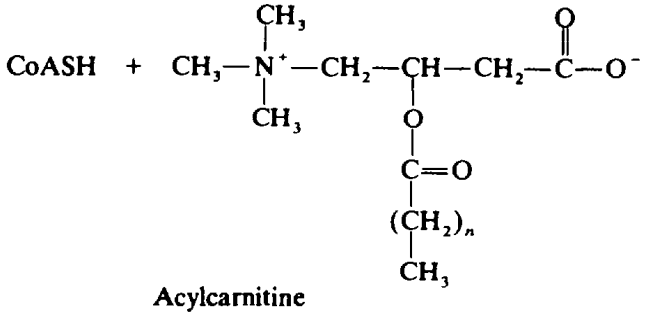
The enzymes that oxidize fatty acids are located in the mitochondrial matrix. Acyl-CoA derivatives do not freely permeate the inner mitochondrial membrane, but a specific transport protein allows entry of the acyl chains to the matrix.

EXAMPLE 13.7

Carnitine is an *acyl-group carrier* that transports fatty acids into and out of the mitochondrial matrix (Fig. 13-4). Acyl groups are linked by esterification to the hydroxyl group of carnitine by the action of *carnitine acyltransferase* that resides in the inner membrane of the mitochondrion.



Acyl-CoA Carnitine



Acylcarnitine

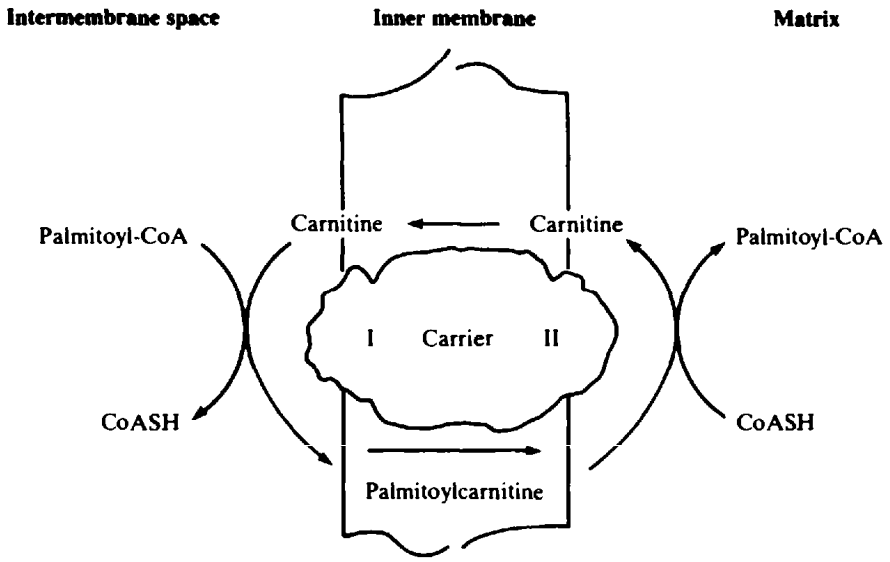


Fig. 13-4 Transport of palmitoyl groups between the cytoplasm and the mitochondrial matrix. I and II are two carnitine palmitoyltransferases. The carriers transport palmitoylcarnitine across the membrane inward, and free carnitine is transported across the membrane outward. Free CoA is not transported.

β-Oxidation of Fatty Acids

β-Oxidation of acyl-CoA derivatives of fatty acids occurs so that fatty acids are sequentially shortened by two carbon units at a time by a process that yields acetyl-CoA as the only product (Fig. 13-5). The acyl chains are cleaved at the bond between C-2 and C-3 of the chain, which is the so-called *β bond*, by a process that induces oxidation of this part of the molecule. Table 13.1 lists the reactions and enzymes for the β-oxidation of fatty acids shown in Fig. 13-5.

Question: What happens to the $FADH_2$ and NADH produced during β oxidation?

Acyl-CoA dehydrogenase is a *flavin-linked*, membrane-bound enzyme, associated with the mitochondrial respiratory complexes. When $FADH_2$ is produced, it is oxidized by the respiratory chain (Chap. 14). Hydroxyacyl-CoA dehydrogenase is in the mitochondrial matrix, and the NADH produced by the action of this enzyme on hydroxyacyl-CoA compounds contributes to the pool of NADH in the matrix. It is also oxidized by the respiratory chain (Chap. 14).

13.6 THE FATE OF ACETYL-CoA FROM FATTY ACIDS: KETOGENESIS

Acetyl-CoA is oxidized to carbon dioxide via the citric acid cycle (Chap. 12), thus transforming additional energy to that which has been transformed via β-oxidation. In liver mitochondria only, acetyl-CoA may also be converted to *ketone bodies*:

CH_3-CO-CH_2-COO^-	CH_3-CO-CH_3	CH_3-CHOH-CH_2-COO^-
Acetoacetate	Acetone	3-Hydroxybutyrate

The reactions and the biosynthetic pathway for these compounds are shown in Table 13.2 and Fig. 13-6.

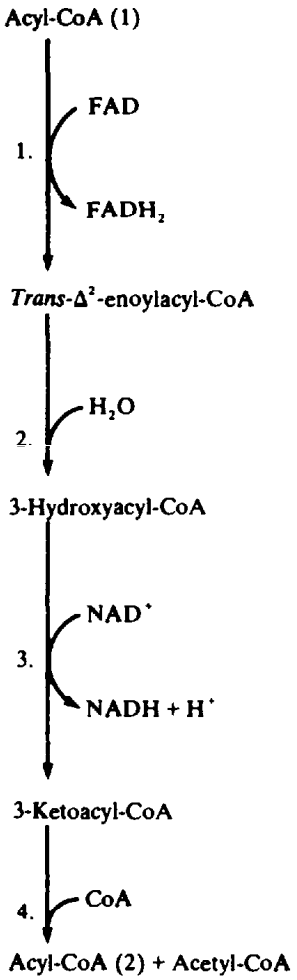


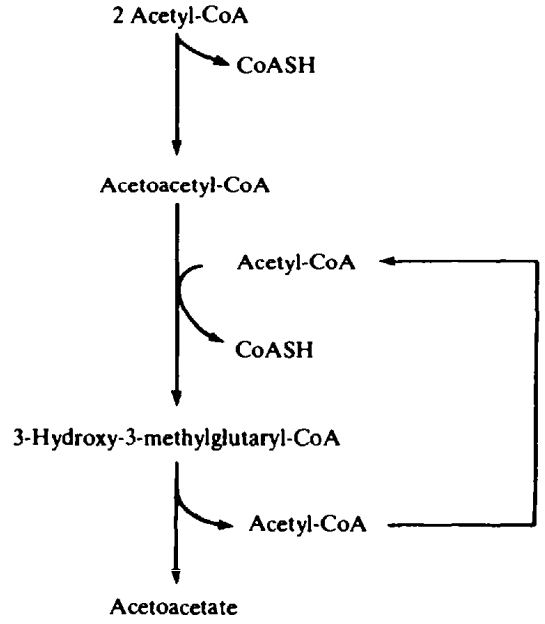
Fig. 13-5 Metabolic pathway of β-oxidation; removal of the first two carbon units of a fatty acid. If the acyl-CoA (1) is palmitoyl-CoA ($C_{16:0}$), then acyl-CoA (2) is myristoyl-CoA ($C_{14:0}$). Hence, the complete β-oxidation of palmitoyl-CoA requires seven such cleavages and produces eight molecules of acetyl-CoA. The numbers to the left of the arrows correspond to the numbered reactions in Table 13.1.

Ketone bodies are water-soluble lipid fuels that are continuously released from the liver. When carbohydrate is plentiful and glucose is readily available to the tissues, the amount of circulating ketone bodies is low ($\sim$0.1 mmol L^{-1}). When large amounts of triacylglycerol are being hydrolyzed in adipose tissues, in response to an increase in whole-body energy demand, the rate of oxidation of fatty acid increases in the liver and other tissues. In the liver, this increases ketogenesis and thus increases the ketone body concentration in the circulation. During exercise, for example, the blood ketone-body concentration rises steadily and can reach 2 to 3 mmol L^{-1}, which is comparable to the blood glucose concentration.

Normally, some acetoacetate is converted to 3-hydroxybutyrate. Acetoacetate and 3-

Table 13.1. The Reactions of β-Oxidation of Fatty Acids

Reaction	Enzyme
1. $CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}H_2\text{—}\overset{②}{C}H_2\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA} + FAD \longrightarrow CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}H\text{=}\overset{②}{C}H\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA} + FADH_2$	Acyl-CoA dehydrogenase
2. $CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}H\text{=}\overset{②}{C}H\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA} + H_2O \longrightarrow CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}HOH\text{—}\overset{②}{C}H_2\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA}$	Enoyl-CoA hydratase
3. $CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}HOH\text{—}\overset{②}{C}H_2\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA} + NAD^+ \longrightarrow CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}O\text{—}\overset{②}{C}H_2\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA} + NADH + H^+$	3-Hydroxyacyl-CoA dehydrogenase
4. $CH_3\text{-}(CH_2)_n\text{-}\overset{③}{C}O\text{—}\overset{②}{C}H_2\text{-}\overset{O}{\overset{\|}{C}}\text{-SCoA} + CoASH \longrightarrow CH_3\text{-}(CH_2)_n\text{-}\overset{③}{\overset{O}{\overset{\|}{C}}}\text{-SCoA} + CH_3\text{-}\overset{②}{\overset{O}{\overset{\|}{C}}}\text{-SCoA}$	Thiolase



3-Hydroxybutyrate and acetone are derived from acetoacetate:

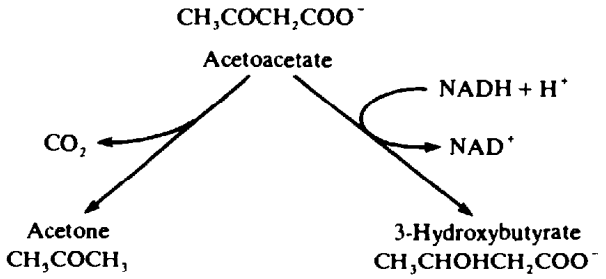


Fig. 13-6 Pathway of ketone body synthesis: ketogenesis. Table 13.2 shows the reactions and enzymes involved in the production of acetoacetate and 3-hydroxybutyrate shown here.

Table 13.2. Ketogenesis

Reaction	Enzyme
1. $2CH_3\text{-}CO\text{-}SCoA \rightarrow CH_3\text{-}CO\text{-}CH_2\text{-}CO\text{-}SCoA$	Acetyl-CoA acetyltransferase
2. $CH_3\text{-}CO\text{-}CH_2\text{-}CO\text{-}SCoA + CH_3\text{-}CO\text{-}SCoA \rightarrow$ $^-OOC\text{-}CH_2\text{-}C(CH_3)(OH)\text{-}CH_2\text{-}CO\text{-}SCoA + CoASH$	Hydroxymethylglutaryl-CoA (HMG-CoA) synthase
3. $^-OOC\text{-}CH_2\text{-}C(CH_3)(OH)\text{-}CH_2\text{-}CO\text{-}SCoA \rightarrow$ $^-OOC\text{-}CH_2\text{-}CO\text{-}CH_3 + CH_3\text{-}CO\text{-}SCoA$	HMG-CoA lyase
4. $CH_3\text{-}CO\text{-}CH_2\text{-}COO^- + NADH + H^+ \rightarrow$ $CH_3\text{-}CHOH\text{-}CH_2\text{-}COO^- + NAD^+$	3-Hydroxybutyrate dehydrogenase

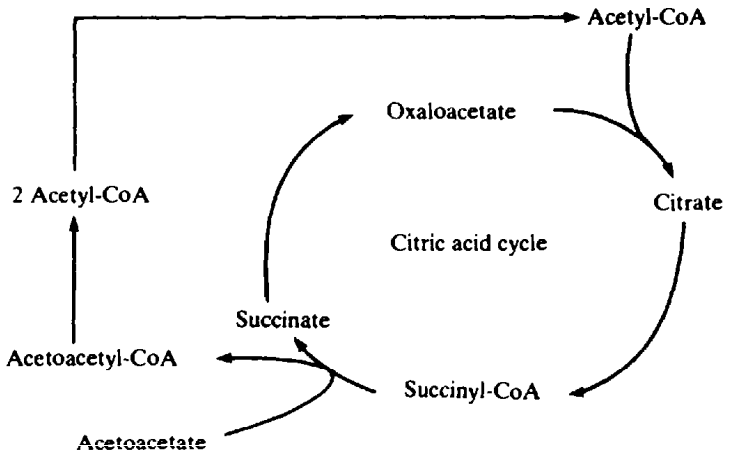


Fig. 13-7 Pathway of acetoacetate utilization.

hydroxybutyrate are valuable fuels for skeletal and cardiac muscle, and it is estimated that they supply 10 percent of the daily energy requirement of these tissues.

Question: How is the energy of acetoacetate made available for mechanical work in muscles?

The enzyme *3-oxoacid transferase* is in mitochondria of muscle and converts acetoacetate to acetoacetyl-CoA; however, it is absent from liver mitochondria. The acetoacetyl-CoA is cleaved to acetyl-CoA by *thiolase*, which is in the mitochondria of all tissues (see Fig. 13-7).

13.7 LIPOGENESIS

When there is an oversupply of dietary carbohydrate, the excess carbohydrate is converted to triacylglycerol; individuals on low-fat diets also convert glucose to triacylglycerol which is stored. This involves the synthesis of fatty acids from acetyl-CoA and the esterification of fatty acids in the production of triacylglycerol. The process is called *lipogenesis*. The major lipogenic tissues are the intestine, liver, and adipose tissue. During lactation, the mammary gland, too, becomes a major site of lipogenesis and places a heavy demand on a continuing supply of glucose for the synthesis of milk lipids.

Synthesis of Palmitic Acid

The synthesis of palmitic acid occurs in the cytosol, from acetyl-CoA. When glucose is abundant and the amount of citrate in the mitochondrial matrix exceeds the demand by the citric acid cycle, the excess citrate is transported out of the mitochondria into the cytosol (Fig. 13-8). Citrate in the cytosol is the source of acetyl groups for fatty acid synthesis, and its metabolism there involves the following enzyme reactions:

(*a*) *Citrate lyase*

$$\begin{array}{l} H_2C\text{-}COO^- \\ | \\ HOC\text{-}COO^- \\ | \\ H_2C\text{-}COO^- \\ \text{Citrate} \end{array} + CoASH + ATP^{4-} \rightarrow \underset{\text{Oxaloacetate}}{^-OOC\text{-}CH_2\text{-}CO\text{-}COO^-} + \underset{\text{Acetyl-CoA}}{CH_3\text{-}CO\text{-}SCoA} + ADP^{3-} + P_i^{2-}$$

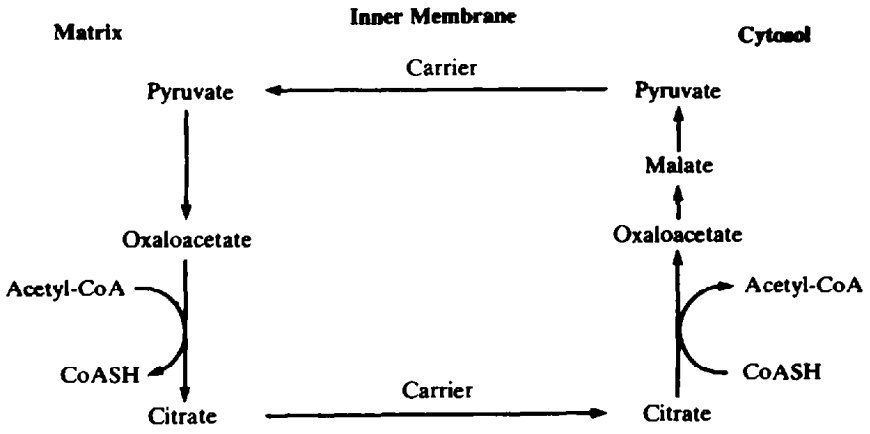


Fig. 13-8 Transfer of acetyl groups from the mitochondrion to the cytosol.

(*b*) *Malate dehydrogenase (1)*

$$^{-}OOC\text{-}CH_2\text{-}CO\text{-}COO^{-} + NADH + H^{+} \rightarrow {}^{-}OOC\text{-}CH_2\text{-}HCOH\text{-}COO^{-} + NAD^{+}$$

Oxaloacetate — Malate

(*c*) *Malate dehydrogenase (2)* (oxaloacetate-decarboxylating)($NADP^{+}$)

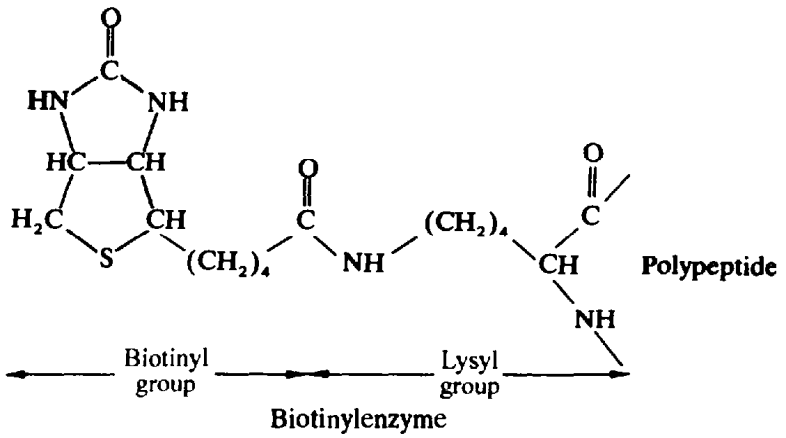
$$^{-}OOC\text{-}CH_2\text{-}CHOH\text{-}COO^{-} + NADP^{+} \rightarrow CH_3\text{-}CO\text{-}COO^{-} + CO_2 + NADPH$$

Malate — Pyruvate

For the conversion of pyruvate to oxaloacetate and the formation of citrate in the mitochondrion, see Chap. 12. Acetyl-CoA for fatty acid synthesis is converted to malonyl-CoA; this reaction is catalyzed by *acetyl-CoA carboxylase*. Seven molecules of acetyl-CoA are converted to malonyl-CoA for the synthesis of one molecule of palmitic acid.

EXAMPLE 13.8

In the formation of malonyl-CoA via acetyl-CoA carboxylase, *biotin* which is tightly bound to the enzyme as a prosthetic group acts as a carrier of a carboxyl group that is transferred to acetyl-CoA.



Normally, acetyl-CoA carboxylase exists in the cytosol as an inactive protomer, $M_r = 2 \times 10^5$. Citrate induces polymerization of the protomer into long filaments of an active, multifunctional enzyme, $M_r = 4.8 \times 10^6$. The

active acetyl-CoA carboxylase activates carbon dioxide and transfers a carboxyl group to acetyl-CoA to form malonyl-CoA:

$$HCO_3^- + ATP + \text{Biotinylenzyme} \rightarrow {}^-OOC\text{-biotinylenzyme} + ADP + P_i$$

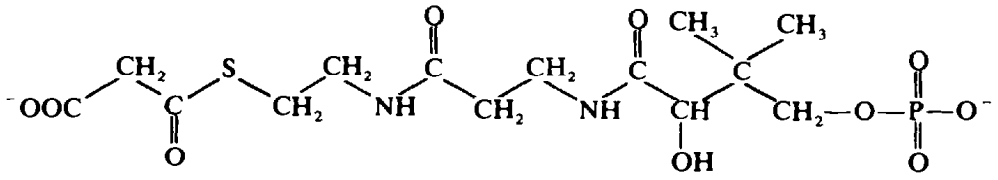
$$CH_3\text{-CO-SCoA} + {}^-OOC\text{-biotinylenzyme} \rightarrow {}^-OOC\text{-}CH_2\text{-CO-SCoA} + \text{Biotinylenzyme}$$

$$\text{Sum: Acetyl-CoA} + HCO_3^- + ATP \rightarrow \text{Malonyl-CoA} + ADP + P_i$$

The formation of malonyl-CoA signals the beginning of the synthesis of palmitic acid ($C_{16:0}$). This occurs on a multifunctional enzyme complex, the *fatty acid synthase*. In mammalian liver, the enzyme complex consists of two identical polypeptides, each with specific binding sites for malonyl and alkanoyl groups, and eight different enzyme activities.

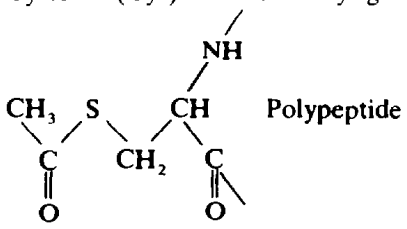
The malonyl and alkanoyl binding sites on the fatty acid synthase are as follows:

(*a*) Phosphopantotheine (PP) binds a malonyl or acetyl group:

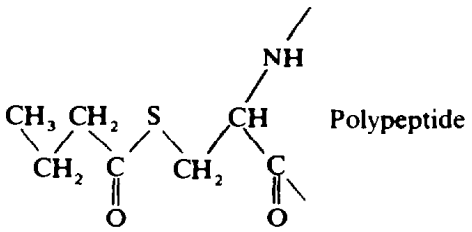


Malonyl-phosphopantotheine

(*b*) Cysteine (Cys) binds alkanoyl groups only:



Acetylcysteinyl-



Butyrylcysteinyl-

Once an acetyl group and a malonyl group are bound to the fatty acid synthase, seven rounds of enzymic reactions proceed for the synthesis of palmitic acid, which is then released from the complex. The overall reaction is:

$$1 \text{ Acetyl-CoA} + 7 \text{ Malonyl-CoA} + 14NADPH + 14H^+ \rightarrow$$
$$\text{Palmitic acid} + 8 \text{ CoASH} + 14NADP^+ + 7H_2O + 7CO_2$$

The individual enzymatic reactions are given in Table 13.3.

Question: What is the source of NADPH for the reduction reactions in palmitic acid synthesis?

NADPH is produced during the transfer of acetyl groups from the mitochondrion, when malate is oxidized to pyruvate and carbon dioxide [see Fig. 13-8 and reaction (*c*) in the accompanying text]. NADPH is also produced when glucose is oxidized and decarboxylated to ribulose 5-phosphate (Chap. 11).

Synthesis of Unsaturated Fatty Acids

Palmitic acid may be converted to stearic acid ($C_{18:0}$) by elongation of the carbon chain. *Desaturation* of stearic acid produces oleic acid ($C_{18:1}\Delta 9$). Linoleic acid ($C_{18:2}\Delta 9,12$), however, cannot be synthesized in mammalian tissues. Therefore, it is an *essential* fatty acid for animals and must be obtained from the diet; it has two important metabolic roles. One is to maintain the fluid state of membrane lipids, lipoproteins, and storage lipids. The other role is as a precursor of *arachidonic* acid, which has a specialized role in the formation of prostaglandins (Sec. 13.9).

Table 13.3. Enzymatic Reactions of Fatty Acid Synthetase

Reaction	Enzyme
1. CH_3-CO-SCoA + HS-PI $\xrightarrow[CoASH]{}$ CH_3-CO-S-PP	Acetyl transacylase
2. CH_3-CO-S-PP + HS-Cys $\rightleftharpoons$ CH_3-CO-S-Cys + HS-PP	Acetyl transacylase
3. $^-$OOC-CH_2-CO-SCoA + HS-PP $\xrightarrow[CoASH]{}$ $^-$OOC-CH_2-CO-S-PP	Malonyl transacylase
4. CH_3-CO-S-Cys + $^-$OOC-CH_2-CO-S-PP $\rightarrow$ HS-Cys + CH_3-CO-CH_2-CO-S-PP + CO_2	β-Ketoacyl synthase
5. CH_3-CO-CH_2-CO-S-PP $\xrightarrow[NADPH+H^+ \quad NADP^+]{}$ CH_3-CHOH-CH_2-CO-S-PP	β-Ketoacyl reductase
6. CH_3-CHOH-CH_2-CO-S-PP $\xrightarrow[H_2O]{}$ CH_3-CH=CH-CO-S-PP	β-Hydroxyacyl dehydratase
7. CH_3-CH=CH-CO-S-PP $\xrightarrow[NADPH+H^+ \quad NADP^+]{}$ CH_3-CH_2-CH_2-CO-S-PP	Enoyl reductase
8. CH_3-CH_2-CH_2-CO-S-PP $\rightarrow$ CH_3-CH_2-CH_2-CO-S-Cys	Acyl transacylase
Return to step 4 for condensation with malonyl-S-PP six times.	
9. CH_3-$(CH_2)_{14}$-CO-S-PP + H_2O $\rightarrow$ CH_3-$(CH_2)_{14}$-COO^- + H^+-HS-Cys	Palmitoyl thioesterase

Question: What are the steps involved in the conversion of linoleic acid to arachidonic acid?

See Fig. 13-9. The carbon chain of linoleic acid is desaturated at position 6. γ-Linolenic acid is elongated by two carbon units, and then another double bond is introduced in the C_{20} chain at position 5.

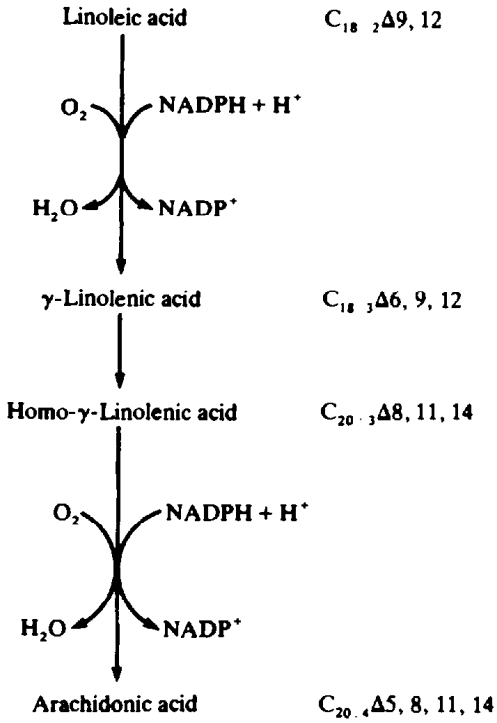
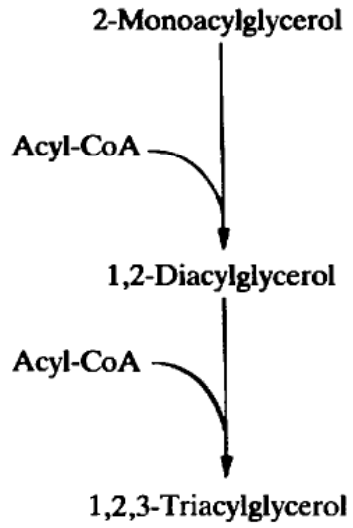


Fig. 13-9 Synthesis of arachidonic acid from linoleic acid.

Triacylglycerol Synthesis

The synthesis of triacylglycerol takes place in the endoplasmic reticulum (ER). In liver and adipose tissue, fatty acids in the cytosol obtained from the diet or from de novo synthesis of palmitic acid become inserted into the ER membrane. The reactions are shown in Fig. 13-10. Membrane-bound *acyl-CoA synthetase* activates two fatty acids, and membrane-bound *acyl-CoA transferase* esterifies them with glycerol 3-phosphate, to form *phosphatidic acid*. *Phosphatidic acid phosphatase* releases phosphate, and in the membrane, 1,2-diacylglycerol is esterified with a third molecule of fatty acid.

In the intestine, triacylglycerol synthesis also occurs in the ER membrane, but fatty acids are esterified with 2-monoacylglycerol, as follows:



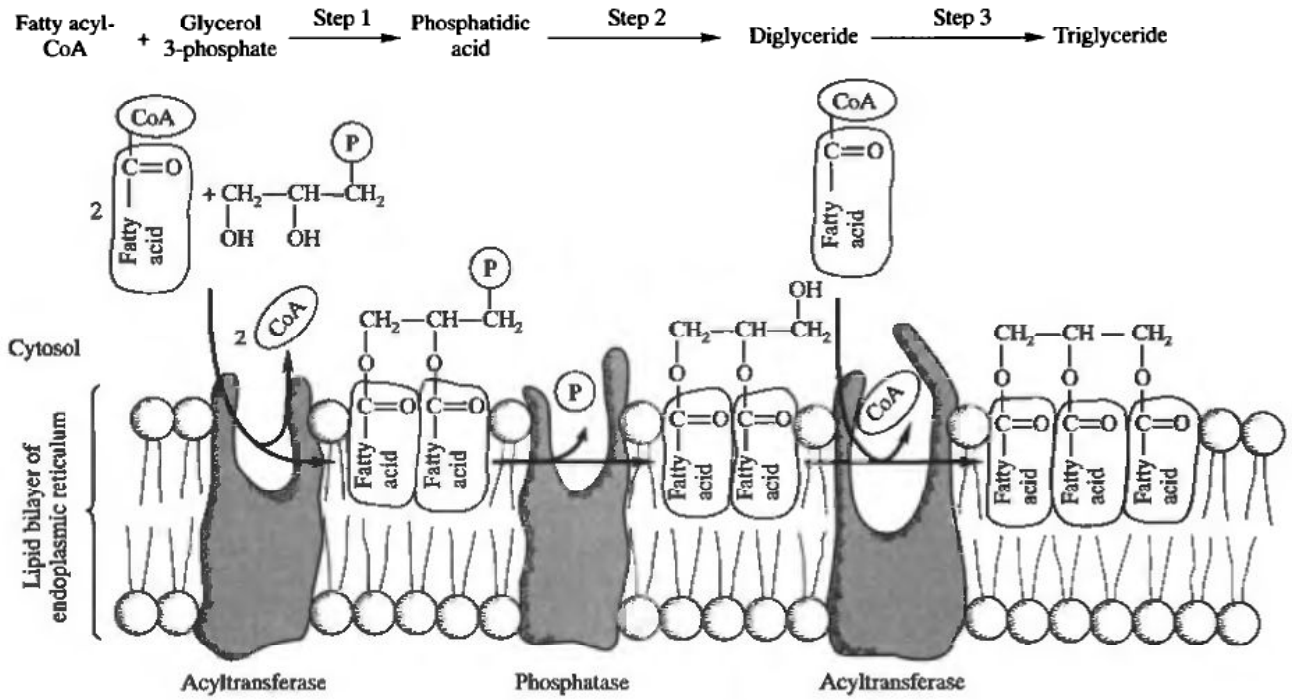


Fig. 13-10 The synthesis of triacylglycerol in the endoplasmic reticulum of liver and fat cells.

Triacylglycerol has no polar interaction with the membrane phospholipids and is either released into the cytosol as tiny lipid droplets or into the lumen of the ER. In fat cells, oil droplets in the cytosol coalesce, migrate toward and fuse with the large central oil droplets. In the liver and intestine, triacylglycerol is packaged into lipoproteins (VLDL and chylomicrons, respectively), which then are secreted into the circulation (see Fig. 13-11).

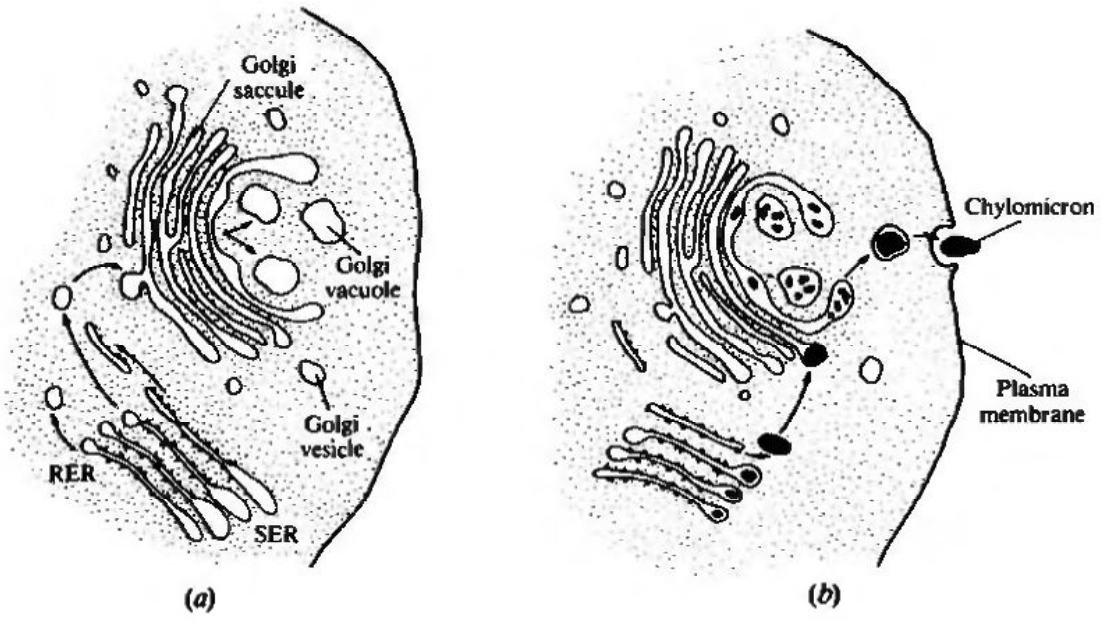


Fig. 13-11 Secretion of chylomicrons and VLDL: (*a*) from rats fasted for 24 h; (*b*) from rats 15–60 min after fat feeding.

13.8 SYNTHESIS OF PHOSPHOLIPIDS AND SPHINGOLIPIDS

Phospholipids

Phosphatidylcholine, a major phospholipid constituent of membranes and lipoproteins, is synthesized de novo in liver cells. The synthesis occurs on the ER and is linked, through 1,2-diacylglycerol, with the synthesis of triacylglycerol. Three compounds specifically involved in the synthesis of phosphatidylcholine are:

1. Choline

$$HO\text{-}CH_2\text{-}CH_2\text{-}N^+(CH_3)_3$$

2. Choline phosphate

$$^{2-}O_3P\text{-}O\text{-}CH_2\text{-}CH_2\text{-}N^+(CH_3)_3$$

3. Cytidine diphosphocholine (CDP-choline)

$$\text{Cytidine-O-}\overset{\overset{O}{\|}}{\underset{\underset{O^-}{|}}{P}}\text{-O-}\overset{\overset{O}{\|}}{\underset{\underset{O^-}{|}}{P}}\text{-O-}CH_2\text{-}CH_2\text{-}N^+(CH_3)_3$$

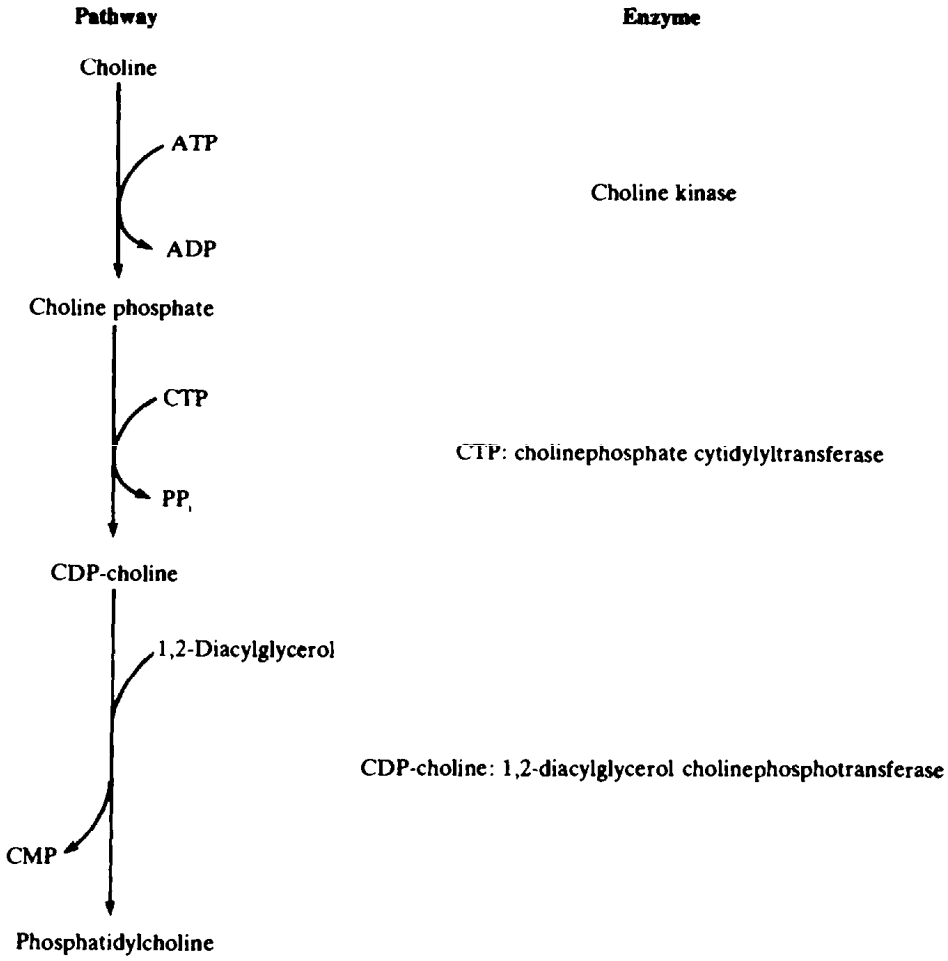


Fig. 13-12 Synthesis of phosphatidylcholine.

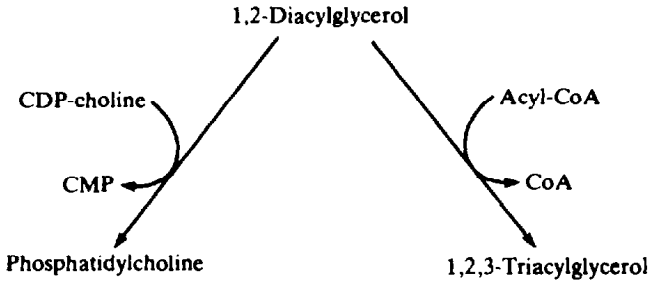


Fig. 13-13 Synthesis of phosphatidylcholine and 1,2,3-triacylglycerol.

The synthesis of phosphatidylcholine is shown in Fig. 13-12.

EXAMPLE 13.9

Note that the enzymes CDP-choline: 1,2-diacylglycerol transferase in phospholipid synthesis and acyl-CoA transferase in triglyceride synthesis have the common substrate 1,2-diacylglycerol; thus we have this two-way synthesis shown in Fig. 13-13.

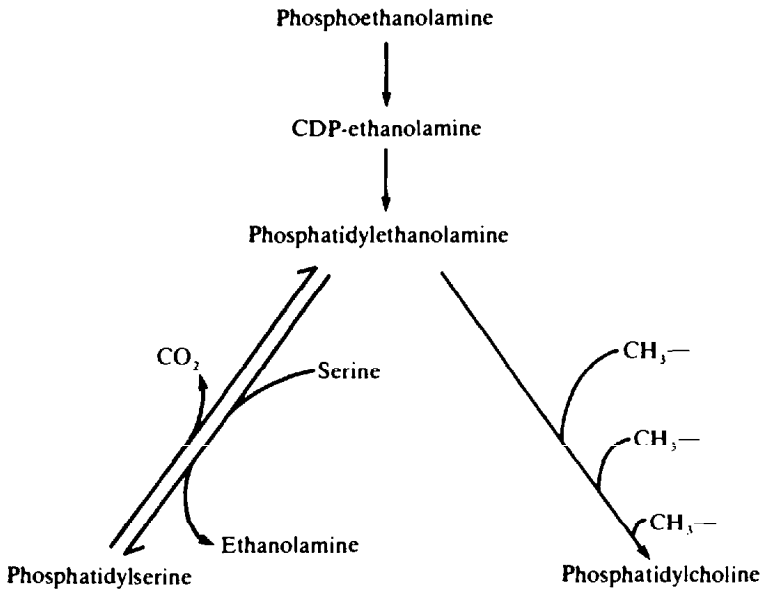


Fig. 13-14 Synthesis of phosphatidylserine and phosphatidylcholine from phosphoethanolamine.

Phosphatidylethanolamine is synthesized de novo in a similar way to phosphatidylcholine. Phosphatidylserine and phosphatidylcholine may then be formed from phosphatidylethanolamine.

EXAMPLE 13.10

Phosphatidylserine arises by an exchange of the ethanolamine residue of phosphatidylethanolamine for a seryl group. Decarboxylation of the serine of phosphatidylserine reforms phosphatidylethanolamine. Three successive methylation reactions convert phosphatidylethanolamine to phosphatidylcholine. *S*-Adenosylmethionine is the methyl-group donor (Chap. 15) (see Fig. 13-14).

Sphingolipids

Sphingolipids comprise *glycolipids* (gangliosides and cerebrosides) and the phospholipid, *sphingomyelin*. These compounds, too, are important membrane constituents. The biosynthesis of sphingolipids involves the common intermediate *ceramide* (Fig. 13-15).

Degradation of Phospholipids and Sphingolipids

During growth and development of tissues, cell material is constantly being degraded and resynthesized. Membrane lipids are degraded by lysosomal enzymes.

Question: What is the sequence of events in the hydrolysis of (*a*) phosphatidylcholine and (*b*) sphingomyelin in lysosomes?

The following reactions can occur in any order:

(*a*) 1. Phospholipase A_1 removes a fatty acid.
2. Phospholipase A_2 removes a fatty acid.
3. Phospholipase C removes choline phosphate.
4. Phospholipase D removes choline.

(*b*) 1. Sphingomyelinase removes choline phosphate from sphingomyelin.
2. Ceramidase removes fatty acid from ceramide.

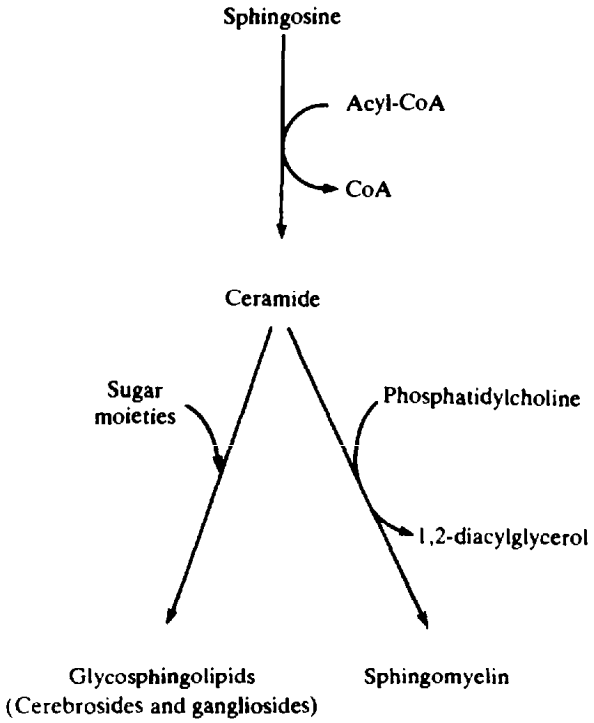


Fig. 13-15 Pathway of synthesis of sphingolipids.

Cerebrosides (ceramide monosaccharides) are important membrane constituents of the brain and central nervous system, and gangliosides (ceramide oligosaccharides that contain sialic acid) have an important role in many tissues as cell-surface receptors for a variety of ligands and hormones. In humans, there are 10 known classes of lipids storage disease, in which the degradation of sphingolipids does not take place adequately. This results in the accumulation of sphingolipid, followed by swelling and malfunction of tissues; it affects the young and usually leads to early death. The most common symptom in all these diseases is mental retardation (see Table 13.4); this emphasizes the importance of breakdown and resynthesis of sphingolipids in nervous tissue. The retina, the liver, and the spleen may also be affected. The diseases are inherited and are due to genetic defects that result in reduced hydrolytic enzyme activity.

Table 13.4. Clinical Signs of Some Lipid-Storage Diseases

Disease	Signs and Symptoms
Gaucher disease	Spleen and liver enlargement Erosion of long bones and pelvis Mental retardation (only in infantile form)
Niemann-Pick disease	Liver and spleen enlargement Mental retardation About 30% with red spot in retina
Metachromatic leukodystrophy	Mental retardation Psychological disturbances in adult form Yellow-brown staining of nerves seen with cresyl violet dye

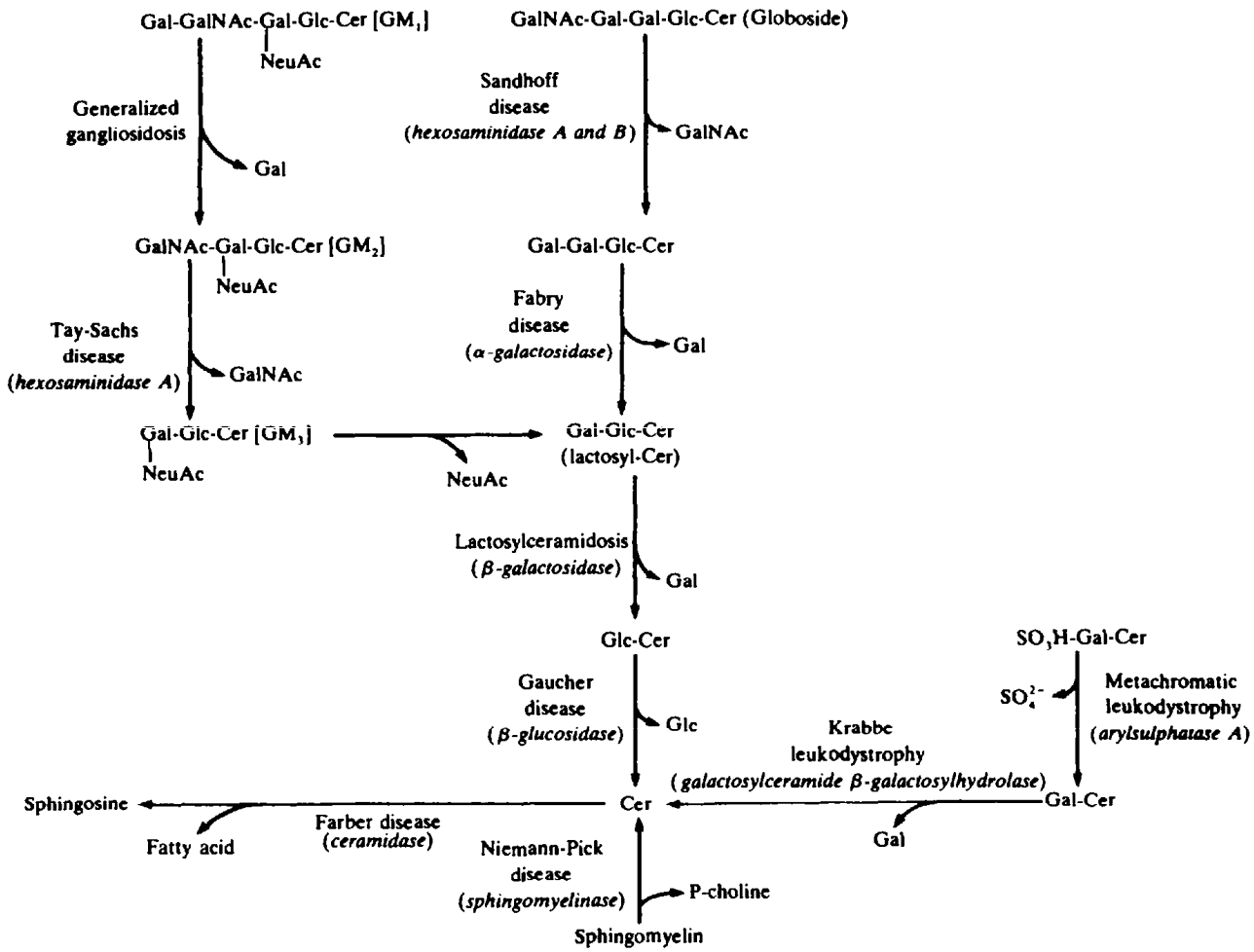


Fig. 13-16 Cer = ceramide; Gal = galactosyl; Glc = glucosyl; GalNAc = *N*-acetylgalactosaminidyl; NeuAc = *N*-acetylneuraminidyl; GM_1, GM_2, and GM_3 = type 1, type 2, and type 3 monosialoganglioside, respectively. The specific enzyme deficiency is listed in parentheses for each disease.

EXAMPLE 13.11

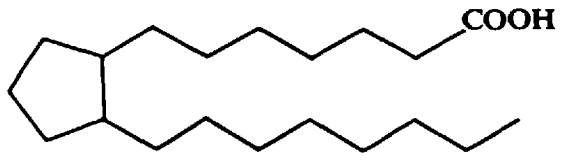
The degradative pathways for the major gangliosides of the ~30 different gangliosides found in humans are shown in Fig. 13-16, together with the common names of the enzymes involved at the various degradation steps, and the names of the diseases that result from a defect in the particular enzyme at each of these steps.

13.9 PROSTAGLANDINS

Structure and Nomenclature

The prostaglandins are C20 unsaturated hydroxy acids with a substituted cyclopentane ring and two aliphatic side chains.

The carbon skeleton of the prostaglandins is:

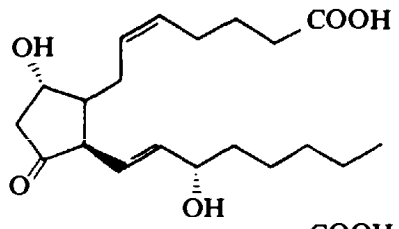


Prostanoic acid

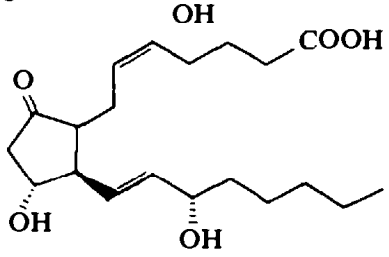
Individual prostaglandins are described by a system of abbreviation in which the name *prostaglandin* is designated PG, followed by a third letter (A–I) that indicates the nature of the substituents on the cyclopentane ring. A numerical subscript indicates the total number of double bonds in the aliphatic chains.

EXAMPLE 13.12

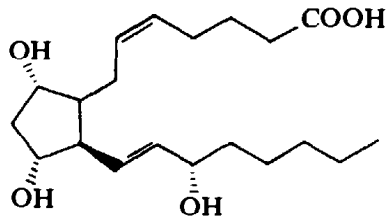
The structures and abbreviated names of some prostaglandins are:



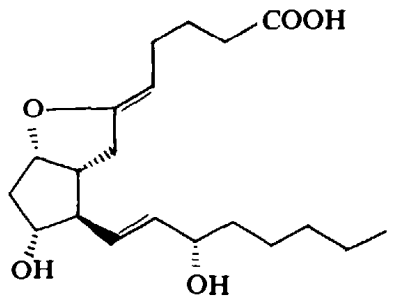
PGD_2



PGE_2



PGF_2



PGI_2
(also called
prostacyclin)

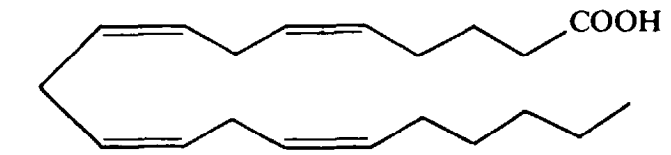
Biological Role of the Prostaglandins

The prostaglandins occur in all tissues but in very small amounts. They act on loci in the same cells as those in which they are synthesized, and their biological roles are diverse; e.g., they function in the female reproductive system during ovulation, menstruation, pregnancy, and parturition, and they stimulate uterine muscle contraction.

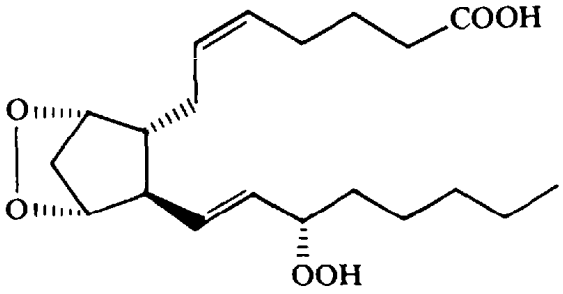
Synthesis of Prostaglandins

Prostaglandins are synthesized as shown in Fig. 13-17 from arachidonic acid in a metabolic pathway that begins with plasma membrane phospholipids. The double-bond arrangement in the carbon chain of arachidonic acid, $C_{20:4}\Delta 5,8,11,14$, makes the fatty acid very susceptible to oxidation

by molecular oxygen. The enzyme *cyclooxygenase* catalyzes the introduction of oxygen and the cyclization of the carbon chain of arachidonic acid in the region of the double-bond positions at C-8 and C-11, as follows:



Arachidonic acid



Prostaglandin G_2 (PGG_2)

Cyclooxygenase is a component of *prostaglandin synthase* that is located in the ER membrane.

The biosynthesis of the primary prostaglandin, PGG_2, leads to the biosynthesis of a large number of chemically related secondary compounds.

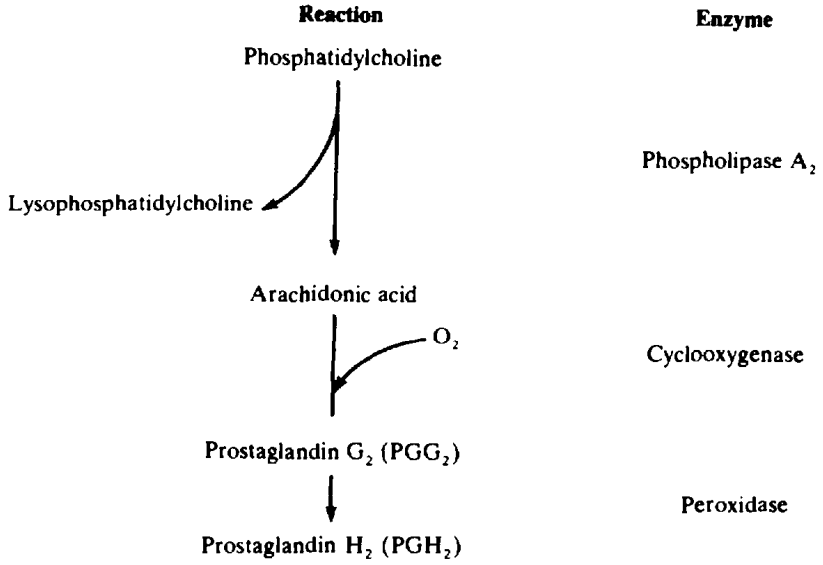


Fig. 13-17

EXAMPLE 13.13

The interrelationships of the prostaglandins are shown in Fig. 13-18.

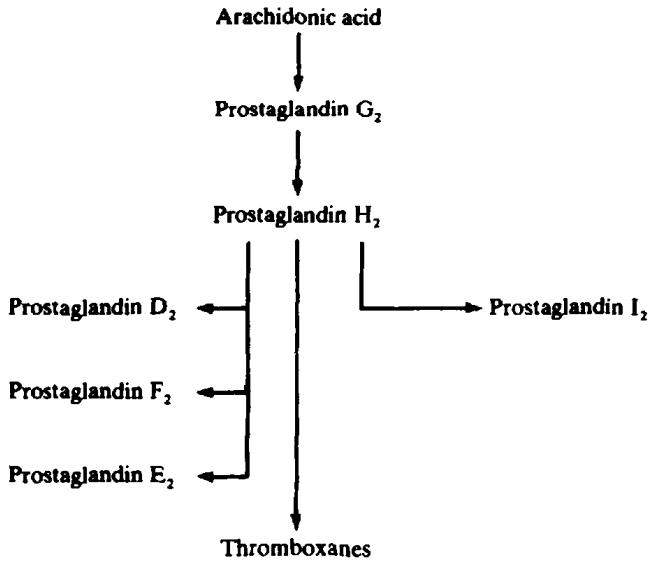
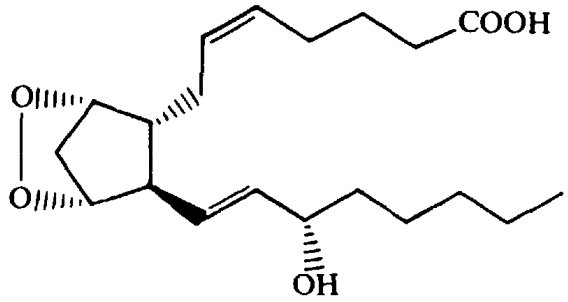


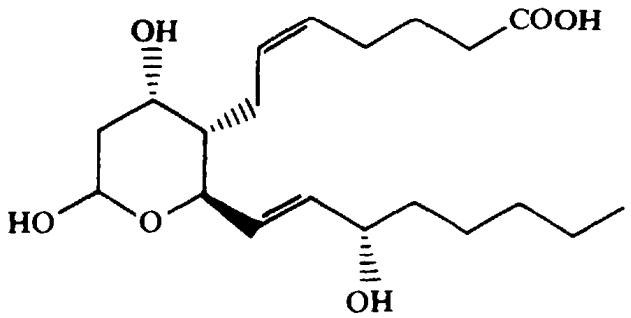
Fig. 13-18

Question: What are the structural characteristics that distinguish prostaglandin H_2 from thromboxanes?

Prostaglandin H_2 (PGH_2) is an *endoperoxide*, the oxygen atoms being attached to the cyclopentane ring:



Thromboxanes have an oxygen atom incorporated into the cyclopentane ring, which produces a six-membered ring:



Thromboxanes were isolated first from platelets and were shown to cause the aggregation of platelets that is involved in the formation of a blood clot.

Leukotrienes

Another group of compounds that are chemically related to the prostaglandins are the *leukotrienes*. These compounds are also derived from arachidonic acid and are linear oxidation products found in leukocytes (white blood cells). Leukotrienes are distinguished by containing a conjugated triene double-bond arrangement. The conversion of arachidonic acid to leukotrienes is as shown in Fig. 13-19.

Leukotriene B_4 is one of the most potent chemotactic agents involved in the inflammatory response in mammalian tissues; a concentration of 1.0 nmol L^{-1} attracts neutrophils to a site of injury or infection.

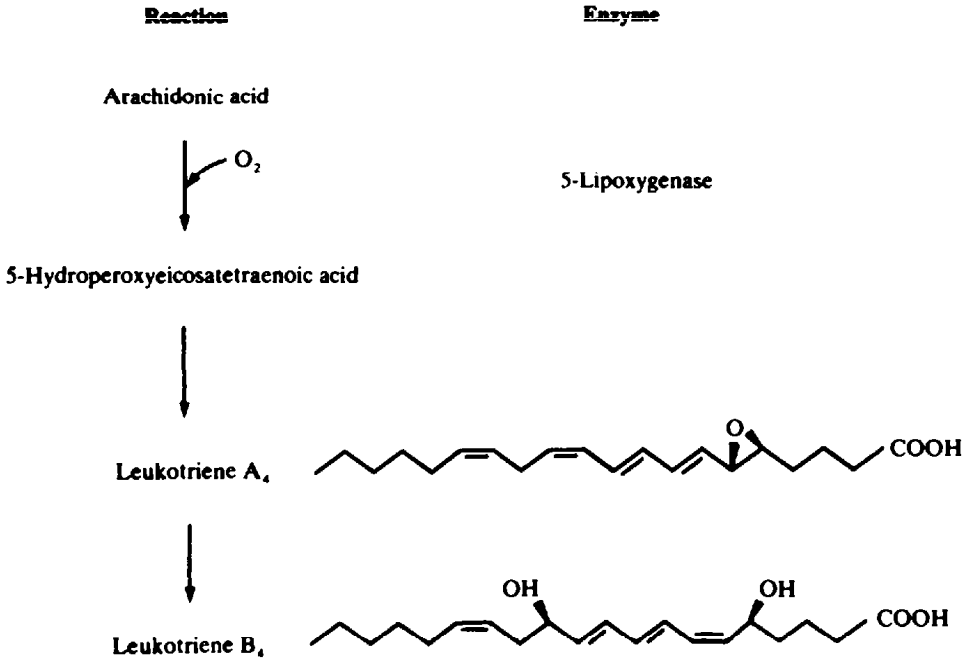


Fig. 13-19 The conversion of arachidonic acid to leukotrienes.

13.10 METABOLISM OF CHOLESTEROL

Biological Role of Cholesterol

Cholesterol is involved in two major biological processes. It is a structural component of cell membranes (Chap. 6) and the parent compound from which *steroid hormones*, *vitamin* D_3 (cholecalciferol), and the *bile salts* are derived. Cholesterol is synthesized de novo in the liver and intestinal epithelial cells and is also derived from dietary lipid. De novo synthesis of cholesterol is regulated by the amount of cholesterol and triglyceride in the dietary lipid.

Cholesterol Biosynthesis in Liver and Intestinal Epithelium

The biosynthesis of cholesterol begins with acetyl-CoA in what is a very complex process involving 32 different enzymes, some of which are soluble in the cytosol and others of which are bound to the ER membrane. The basic carbon building block of cholesterol is *isoprene* (Chap. 6).

The key intermediates in the biosynthesis of cholesterol are shown in Fig. 13-20.

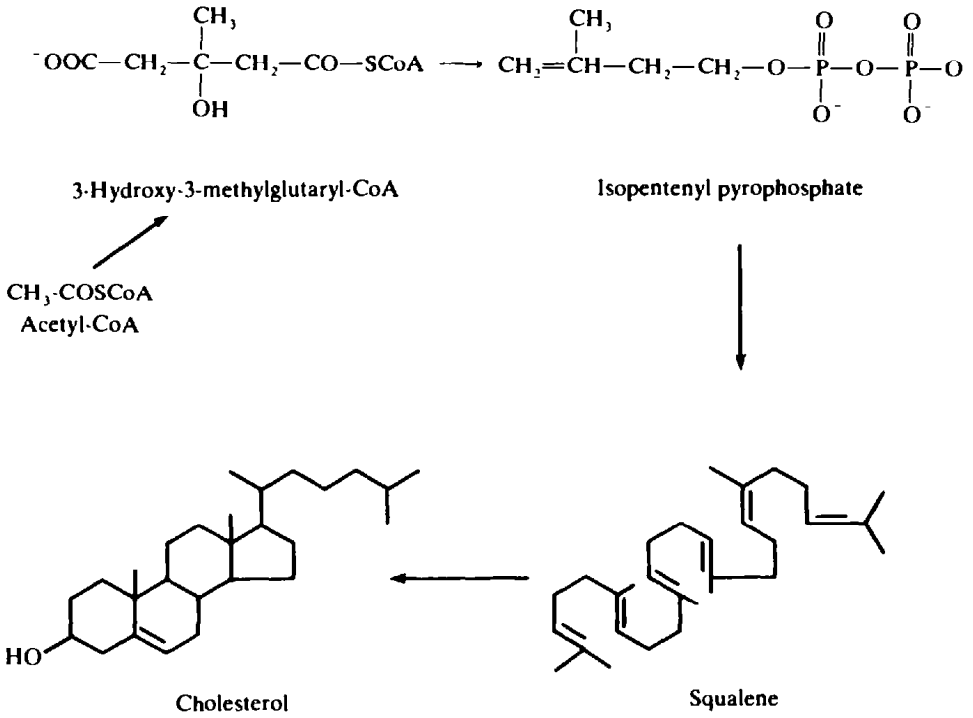


Fig. 13-20 Cholesterol biosynthesis.

In the initial reactions of the biosynthetic pathway, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) is formed from acetyl-CoA by the action of *thiolase* and *HMG-CoA synthase* in the cytosol of the liver cell.

EXAMPLE 13.14

The reactions catalyzed by thiolase and HMG-CoA synthase are:

(*a*) Thiolase

$$2CH_3CO{-}SCoA \longrightarrow CH_3COCH_2CO{-}SCoA + CoASH$$

Acetyl-CoA Acetoacetyl-CoA

(*b*) HMG-CoA synthase

$$H_2O + CH_3COCH_2CO{-}SCoA + CH_3CO{-}SCoA \longrightarrow {}^{-}OOC\text{-}CH_2\text{-}\underset{\displaystyle OH}{\overset{\displaystyle CH_3}{\overset{|}{\underset{|}{C}}}}\text{-}CH_2\text{-}CO{-}SCoA + H^+ + CoASH$$

Acetoacetyl-CoA Acetyl-CoA 3-Hydroxy-3-methylglutaryl-CoA

Question: By what biochemical means is 3-hydroxy-3-methylglutaryl-CoA reduced to the isoprene compound, isopentenyl pyrophosphate?

NADPH is the electron donor in a reaction catalyzed by *hydroxymethylglutaryl-CoA reductase*, followed by two phosphorylation reactions involving ATP and two specific kinases, and an ATP-dependent decarboxylation. The process is summarized in Example 13.15 and Fig. 13-21.

EXAMPLE 13.15

The formation of isopentenyl pyrophosphate is shown in Fig. 13-21.

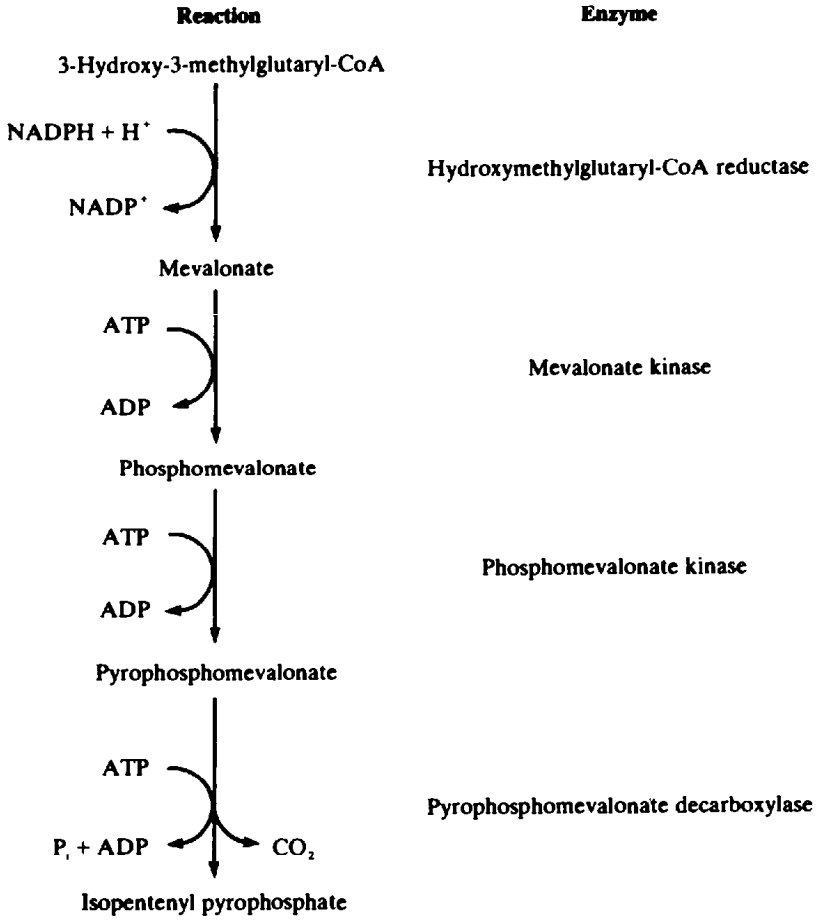


Fig. 13-21 The formation of isopentyl pyrophosphate.

Some isopentenyl pyrophosphate is converted to the isomer *dimethylallyl pyrophosphate*, by an isomerase that produces a mixture, *isopentenyl pyrophosphate* $\rightleftharpoons$ *dimethylallyl pyrophosphate*. From this point on, the carbon chain-length of the intermediates progressively increase; this is followed by reduction to *squalene*, which has 30 carbon atoms in a folded chain and no oxygen atoms (Chap. 6). The conversion of isopentenyl pyrophosphate to squalene is shown in Fig. 13-22.

The final stage of the biosynthesis of cholesterol requires molecular oxygen, and the chain of squalene is cyclized to produce the primary steroid lanosterol, which is subsequently modified to cholesterol. The cyclization of squalene is shown in Fig. 13-23.

Question: How is the biosynthesis of cholesterol regulated by the amount of cholesterol in the diet?

A feedback mechanism operates in which intracellular free cholesterol inhibits *HMG-CoA reductase*. When the diet is rich in cholesterol, intracellular cholesterol increases in the liver and the biosynthesis of cholesterol is suppressed. Conversely, a low-cholesterol diet, but one with adequate triglyceride, stimulates cholesterol biosynthesis.

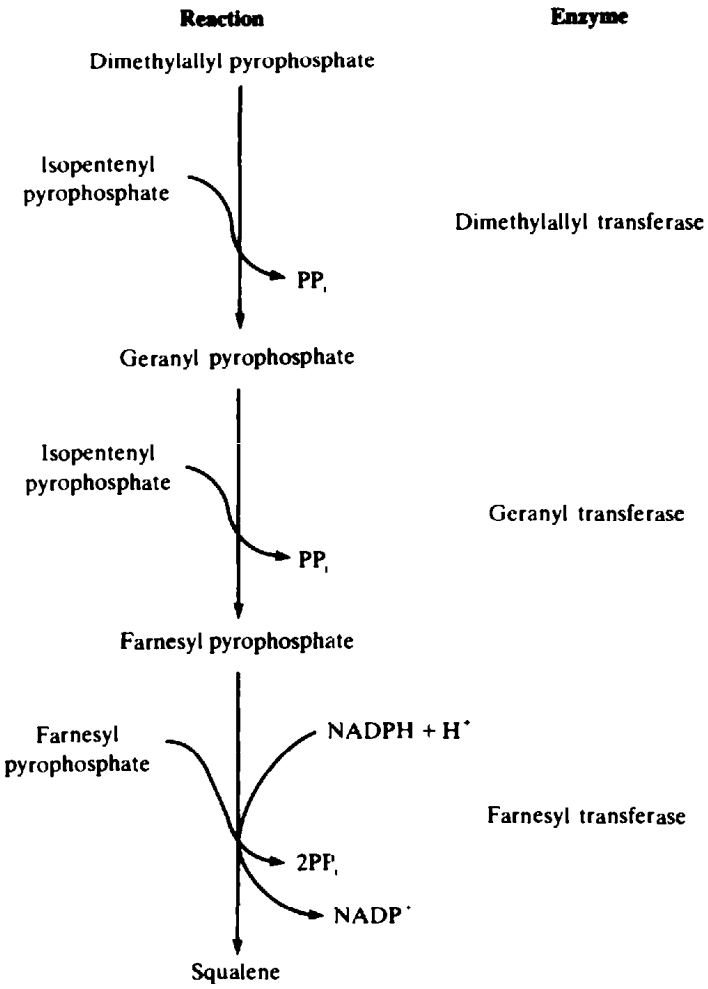


Fig. 13-22 The formation of squalene.

The metabolism of cholesterol involves modifications to the alkyl side chain and the introduction of substituents onto the phenanthrene ring. A group of enzymes, *monooxygenases*, plays an important role in catalyzing the formation of bile salts in the liver and steroid hormones in the adrenal gland and gonads.

Regulation of Cholesterol Synthesis

Cholesteryl esters that are internalized via the LDL receptor are hydrolyzed to produce cholesterol and an acyl chain. Cholesterol, in turn, activates the enzyme *acyl-CoA cholesterol acyltransferase* (ACAT) which re-esterifies cholesterol. In an apparently futile cycle, the cholesteryl esters are hydrolyzed by *cholesteryl ester hydrolase*. The cholesterol moiety has several fates: it may leave the cell and bind to an acceptor such as high-density lipoprotein (HDL), it may be converted to steroid hormones, or it may be reesterified by ACAT. When the cellular cholesterol concentration falls, the activity of HMG-CoA reductase is increased, as is the number of LDL receptors, which results in an increase of cellular cholesterol, due both to *de novo* synthesis and to the uptake of cholesterol-rich lipoproteins in the circulation. An increase in cellular cholesterol results in the rapid decline in the mRNA levels for both HMG-CoA reductase and the LDL receptor. This coordinated regulation is brought about by the presence of an eight nucleotide sequence on the genes which code for both proteins; this is termed the *sterol regulatory element-1*.

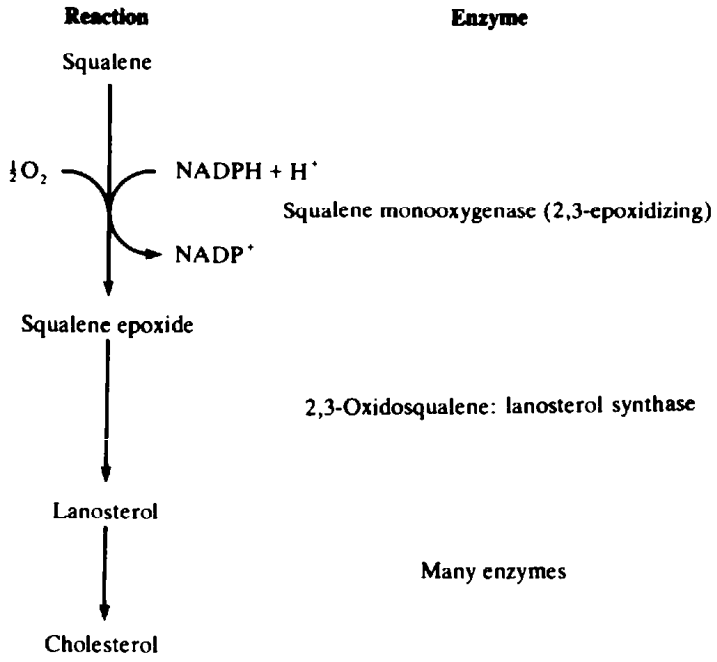


Fig. 13-23 The formation of cholesterol from squalene.

The Bile Salts

The digestion and absorption of dietary lipid can be completed only in the presence of adequate amounts of *bile salts* that are synthesized in the liver and pass, via the bile duct, into the duodenum and thence into the jejunum. Reabsorption of the bile salt micelles occurs in the ileum, from which a large proportion return via the blood to the liver. The bile ducts carry bile salts from the liver to the gallbladder, where they are stored; excreted (excess) cholesterol is dissolved in the bile salt micelles. Overall, 90 percent of the bile salts involved in absorption of lipid in the jejunum are recycled, in a process called the *enterohepatic circulation*, and 10 percent are lost in the feces. Replacement of this amount necessitates conversion from cholesterol. Thus, de novo synthesis of cholesterol itself plays an important part in maintaining the supply of bile salts.

The conversion of cholesterol to bile salts begins when hydroxyl groups are introduced into the phenanthrene ring of cholesterol by the action of cholesterol 7-α-hydroxylase, followed by modification of the side chain. *Cholic acid* and *chenodeoxycholic acid* are produced, as shown in Fig. 13-24.

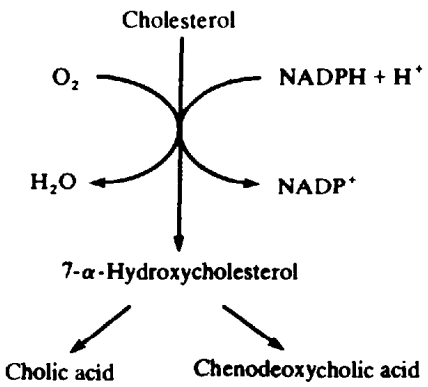


Fig. 13-24 Production of cholic acid and chenodeoxycholic acid.

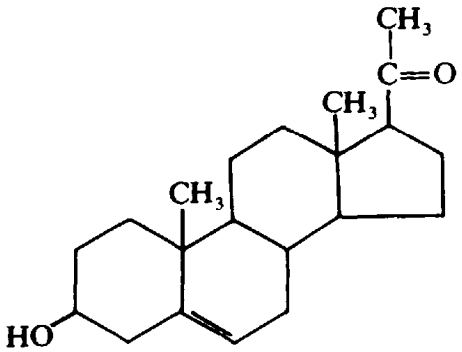
The rate-limiting step of bile acid synthesis is cholesterol 7-α-hydroxylase. The changes in enzyme activity are the result of altered levels of cholesterol 7-α-hydroxylase mRNA.

Question: Why do monooxygenase reactions require NADPH as a cosubstrate?

Monooxygenase reactions catalyze the introduction of only one of the two oxygen atoms from molecular oxygen to form a hydroxyl or keto group in the substrate. The other oxygen atom ends up in water. Both the substrate and the NADPH act as proton and electron donors. Monooxygenase reactions occur in the ER membrane and involve iron-sulfur proteins, ferredoxin, and cytochrome P_{450}.

Steroid Hormones

The synthesis of all steroid hormones begins with the conversion of cholesterol to *pregnenolone*. The pathway of this conversion is shown in Fig. 13-25. The side chain of cholesterol is cleaved by three successive monooxygenase reactions, which introduce a keto group at the site of cleavage of the side chain.



Pregnenolone

Subsequent molecular changes to pregnenolone give rise to the other steroid hormones. All these changes, catalyzed by monooxygenases, involve the introduction of oxygen atoms as either hydroxyl or keto groups at specific sites on the phenanthrene ring of the sterol, and further removal of the side chain (Fig. 13-26).

13.11 REGULATION OF LIPID METABOLISM

The release of fatty acids from adipose tissue is regulated by the rate of hydrolysis of triacylglycerol and the rate of esterification of acyl-CoA with glycerol 3-phosphate. The rate of hydrolysis is stimulated by hormones that bind to cell-surface receptors and stimulate *adenylate cyclase* (which catalyzes the production of cAMP from ATP). *Hormone-sensitive lipase* (Sec. 13.4) can exist in two forms, one of which exhibits very low activity and a second which is phosphorylated and has high activity. Before hormonal stimulation of adenylate cyclase, the low-activity lipase predominates in the fat cell. Stimulation of protein kinase by an increase in cAMP concentration leads to phosphorylation of the low-activity lipase. An increase in the rate of hydrolysis of triacylglycerol and the release of fatty acids from the fat cell follows. This leads to a greater utilization of fatty acids by tissues such as heart, skeletal muscle, and liver.

In liver, β-oxidation and reesterification of acyl-CoA are both possible. The rate of β-oxidation is determined initially by the rate at which acyl groups enter the mitochondrial matrix. This rate of entry may be decreased by *malonyl-CoA*, which inhibits *carnitine palmitoyltransferase* (an enzyme

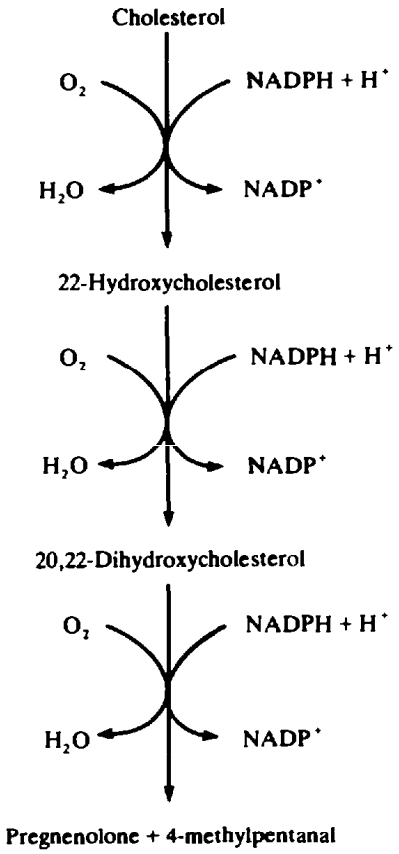


Fig. 13-25 Synthesis of pregnenolone.

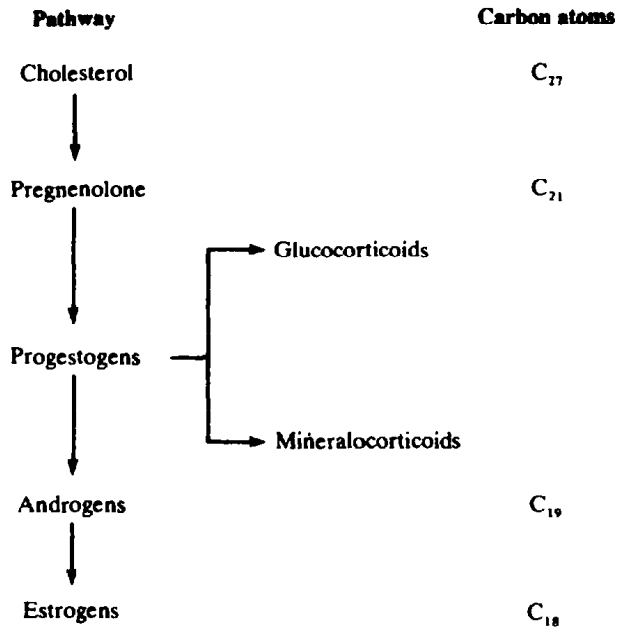


Fig. 13-26 The interrelationships between cholesterol and steroid hormones.

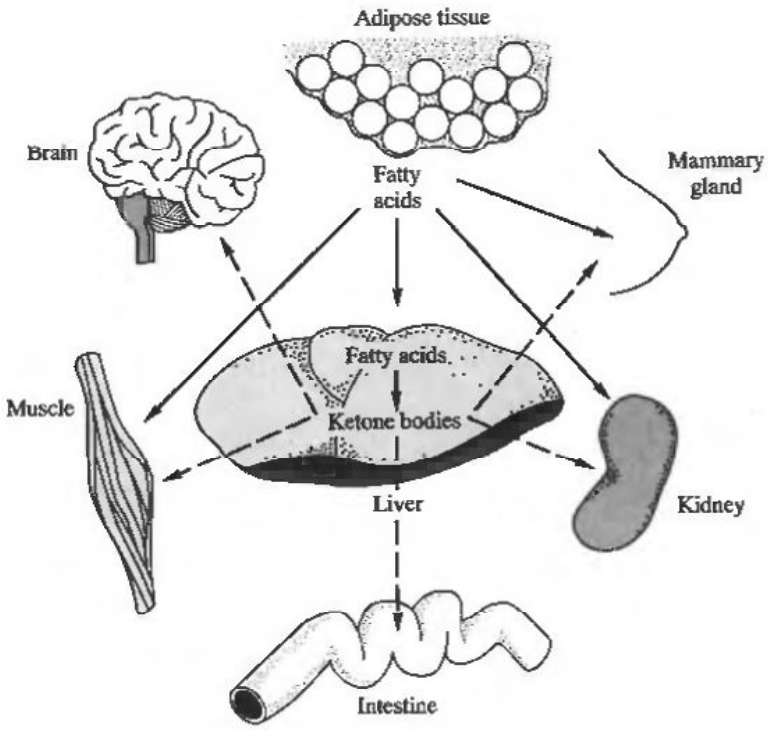


Fig. 13-27 Utilization of ketone bodies during starvation.

that enables acyl group entry to the matrix). During lipogenesis, the concentration of malonyl-CoA in the cytosol is sufficient to inhibit the transferase and thus maintain esterification of fatty acids. During starvation, the release of fatty acids from adipose tissues increases; the rate of ketogenesis in liver also increases. During a 1- to 24-day period of starvation, the concentration of ketone bodies in the blood increases to ~8 mmol L^{-1}. The brain adapts to utilize ketone bodies as fuel; 70 percent of its energy requirement is satisfied by them, and the remainder by glucose. The whole-body utilization of glucose decreases as a number of tissues utilize ketone bodies (Fig. 13-27). In very long starvation periods (greater than 24 days), 3-hydroxybutyrate has a regulatory effect on the release of fatty acids from adipose tissues, probably by increasing the sensitivity of these tissues to insulin.

Insulin is an *antilipolytic* hormone, and its effect on adipose tissue is to increase the transport of glucose into the fat cell, to stimulate lipogenesis and inhibit lipolysis. Thus, pyruvate dehydrogenase and acetyl-CoA carboxylase are activated, and the hormone-sensitive lipase is *in*activated. In the normal, well-fed state insulin stimulates the deposition of fat.

Insulin also exerts a stimulatory effect on the synthesis of cholesterol in the liver. In this tissue, HMG-CoA reductase is activated. HMG-CoA reductase, like hormone-sensitive lipase, can exist in two forms; one is phosphorylated (inactive) and the other is dephosphorylated (active). Phosphorylation of the enzyme depends on an increase in the cellular concentration of cAMP and activation of protein kinase. The dephosphorylation (activation) is catalyzed by a phosphatase. In fat cells, a similar phosphatase dephosphorylates (inactivates) hormone-sensitive lipase. Insulin stimulates the activity of the phosphatase in both liver and fat cells. In this way, active HMG-CoA reductase predominates in the liver cell and directs HMG-CoA into cholesterol synthesis, and in the fat cell hormone-sensitive lipase is *in*activated.

Solved Problems

LIPID DIGESTION

13.1. What are the consequences of abnormal deconjugation of bile salts by bacteria in the small intestine?

SOLUTION

The pH of the contents of the lumen of the small intestine is between 6.0 and 8.0. Bile salts in the lumen are, thus, ionized: e.g., taurocholate has a $pK_a \simeq 1.5$, owing to the conjugation of taurine with the cholate, and glycocholate has a $pK_a \simeq 3.7$, owing to glycine conjugation. Deconjugation of either bile salt leaves free cholate, $pK_a = 5.0$. The higher pK_a value of cholate reduces its solubility in the aqueous environment of the lumen, compared with the bile salts, which are readily soluble. Reabsorption of bile salts in the ileum is thus decreased, and bile acids are excreted.

OXIDATION OF FATTY ACIDS

13.2. How much energy, in the form of ATP, is obtained from β-oxidation of 1 mole of palmitoyl-CoA?

SOLUTION

One mole of palmitoyl-CoA yields eight moles of acetyl-CoA by β-oxidation. The overall equation is:

$$\text{Palmitoyl-CoA} + 7\text{FAD} + 7\text{NAD}^+ + 7\text{CoA} + 7\text{H}_2\text{O} \rightarrow 8\text{Acetyl-CoA} + 7\text{FADH}_2 + 7\text{NADH} + 7\text{H}^+$$

$FADH_2$ and $NADH + H^+$ are oxidized in the electron-transport assemblies of the mitochondria:

$$7\text{FADH}_2 + 7\text{NADH} + 7\text{H}^+ + 7\text{O}_2 \rightarrow 7\text{FAD} + 7\text{NAD}^+ + 14\text{H}_2\text{O}$$

During electron transport, coupled with ATP synthesis, each mole of $FADH_2$ that is oxidized yields 2 moles of ATP; thus, 7 moles of $FADH_2$ yields 14 moles of ATP. Each NADH mole that is oxidized yields 3 moles of ATP; therefore, 7 moles of NADH + 7 moles of H^+ yields 21 moles of ATP. Therefore, the total yield is 35 moles of ATP.

13.3. What is the sequence of events in the complete oxidation of linoleic acid?

SOLUTION

β-Oxidation proceeds normally to shorten the 18-carbon chain to dodecadienoic acid, $C_{12:2}\Delta 3,6$; three molecules of acetyl-CoA are produced as follows:

$$\text{CH}_3\text{-(CH}_2)_4\text{-CH=CH-CH}_2\text{-CH=CH-(CH}_2)_7\text{-COO}^- \rightarrow$$
$$\text{CH}_3\text{-(CH}_2)_4\text{-CH=CH-CH}_2\text{-CH=CH-CH}_2\text{-COO}^- + 3\text{CH}_3\text{-COSCoA}$$

Before the next round of β-oxidation can proceed, the double bond *cis* Δ^3 is converted to *trans* Δ^2 by the enzyme Δ^3-cis-Δ^2-trans-*enoyl-CoA isomerase* as follows:

$$\text{CH}_3\text{-(CH}_2)_4\text{-CH=CH-CH}_2\text{-CH}\underset{cis}{=}\text{CH-CH}_2\text{-COO}^- \rightarrow$$
$$\text{CH}_3\text{-(CH}_2)_4\text{-CH=CH-(CH}_2)_2\text{-CH}\underset{trans}{=}\text{CH-COO}^-$$

After one further β-oxidation cycle, a 4-*cis*-enoyl CoA intermediate is formed. It is acted upon by *enoyl-CoA dehydrogenase* to give 2-*trans*, 4-*cis*-dienoyl CoA. Further metabolism of this intermediate proceeds through one cycle of β-oxidation and requires a second *auxiliary* enzyme, *2,4-dienoyl-CoA reductase* which has high activity in mitochondria. Thus, nine molecules of acetyl-CoA are produced from the oxidation of linoleic acid.

THE FATE OF ACETYL-CoA FROM FATTY ACIDS: KETOGENESIS

13.4. What determines the relative concentrations of acetoacetate and 3-hydroxybutyrate in the blood?

SOLUTION

Acetoacetate is converted to 3-hydroxybutyrate by an *NAD-dependent dehydrogenase*. The reaction is reversible, and the equilibrium constant is approximately 1.

$$K_e = \frac{[\text{3-hydroxybutyrate}][\text{NAD}^+]}{[\text{acetoacetate}][\text{NADH}][\text{H}^+]} \approx 1$$

13.5. 3-Hydroxybutyrate is a higher-energy-content fuel than acetoacetate. How is this energy released in cardiac and skeletal muscle?

SOLUTION

3-Hydroxybutyrate is activated to 3-hydroxybutyryl-CoA and oxidized via a specific 3-hydroxybutyryl-CoA dehydrogenase as in the β-oxidation pathway to acetoacetyl-CoA. This is then converted to acetyl-CoA, which is oxidized in the citric acid cycle.

LIPOGENESIS

13.6. How much glucose is converted to ribulose 5-phosphate when one molecule of palmitic acid is synthesized from acetyl-CoA?

SOLUTION

When glucose is converted to ribulose 5-phosphate, NADPH is produced in the redox reactions of the oxidative part of the pentose phosphate pathway. For every molecule of glucose entering the pathway, two $NADP^+$ molecules are reduced. Some of this NADPH is utilized in the synthesis of palmitic acid. To calculate the amount consider the following:

1. Seven rounds of reactions occur on *fatty acid synthase* to synthesize one molecule of palmitic acid.
2. In each round of the synthase there are two reduction reactions that use NADPH as the electron donor. Therefore, the synthesis of one molecule of palmitic acid requires 14 molecules of NADPH.
3. Eight acetyl groups are transferred from the mitochondrion to the cytosol to provide the carbon for the synthesis of one molecule of palmitic acid. During the transfer of the acetyl groups, 8 molecules of malate are oxidatively decarboxylated to 8 pyruvate and 8 carbon dioxide molecules, and 8 $NADP^+$ molecules are reduced. Thus, 8 of the 14 molecules of NADPH required are provided during the transfer of acetyl groups; the remaining 6 NADPH molecules come from the pentose phosphate pathway from three molecules of glucose.

13.7. What are the characteristics of the fatty acid synthase in the bacterium *Escherichia coli* that distinguish it from the mammalian multienzyme complex?

SOLUTION

Each of the enzymatic activities located in a single polypeptide chain of the mammalian fatty acid synthetase exists as a distinct protein in *E. coli*. The acyl-carrier protein (ACP) of *E. coli* has an $M_r = 8{,}847$ and contains 4-phosphopantotheine. The dehydratase has a molecular weight of 28,000 and catalyzes either trans 2–3 or cis 3–4 dehydration of the hydroxy acid intermediates in the biosynthesis of palmitic acid. When the chain length of the hydroxy fatty acid is C_{10} the synthesis of palmitoleic acid is achieved as follows:

$$\underset{\text{3-Hydroxydecanoyl-ACP}}{CH_3\text{-}(CH_2)_6\text{-}CHOH\text{-}CH_2\text{-}CO\text{-}ACP} \rightarrow \underset{\Delta^3\text{-Decenoyl-ACP}}{CH_3\text{-}(CH_2)_5\text{-}CH{=}CH\text{-}CH_2\text{-}CO\text{-}ACP}$$

Three more rounds of normal synthase activity elongates Δ^3-Decenoyl-ACP to yield

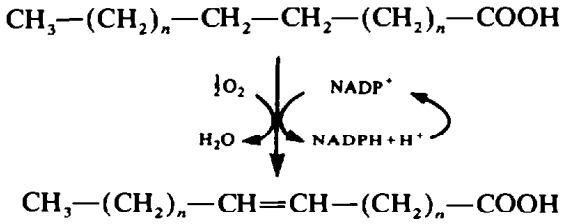
$$CH_3\text{-}(CH_2)_5\text{-}CH{=}CH(CH_2)_7COOH$$

Palmitoleic acid

13.8. (*a*) What is the enzymatic reaction that converts saturated fatty acids to their monounsaturated derivatives in mammalian tissues, and (*b*) in which cellular location does the reaction take place?

SOLUTION

(*a*) An enzyme complex consisting of *acyl-CoA desaturase* and *cytochrome* b_5 catalyzes the oxidation of the alkanoyl chain as follows:



(*b*) This electron-transport chain resides in the ER.

SYNTHESIS OF PHOSPHOLIPIDS AND SPHINGOLIPIDS

13.9. Why does a dietary deficiency of choline in humans and laboratory animals induce a *fatty* liver, i.e., a liver in which the hepatocytes contain large lipid droplets?

SOLUTION

The de novo synthesis of phosphatidylcholine requires choline. In liver, phosphatidylcholine is synthesized and enters the membranes and lipoproteins. The de novo synthesis of phosphatidylcholine also requires 1,2-diacylglycerol. If there is insufficient choline available, phosphatidylcholine production by the de novo pathway cannot occur. The 1,2-diacylglycerol is then converted to triacylglycerol, which accumulates, as it is not secreted in lipoproteins, by the liver. Therefore, the liver cells fill with triacylglycerol.

13.10. How does chronic ingestion of ethanol lead to the development of a fatty liver, if you are told that ethanol stimulates the activity of phosphatidic acid phosphatase?

SOLUTION

The stimulation of phosphatidic acid phosphatase by ethanol stimulates the production of diacylglycerol which in turn stimulates the synthesis of triacylglycerol. Therefore, the triacylglycerol concentration increases in the liver cells.

13.11. Tay-Sachs disease, the lipid storage disease that afflicts 1 in 900 children of Ashkenazic Jewish parents, results in the accumulation of the ganglioside ceramide-β-glucose-β-galactose-*N*-acetylgalactosamine *N*-acetylneuramic acid in the brain. Why does this occur?

SOLUTION

Brain lysosomes are deficient in the enzyme hexosaminidase A. The enzyme has two subunits, A and B. Because of a lack of the enzyme, hydrolysis of the terminal N-acetylgalactosamine from the

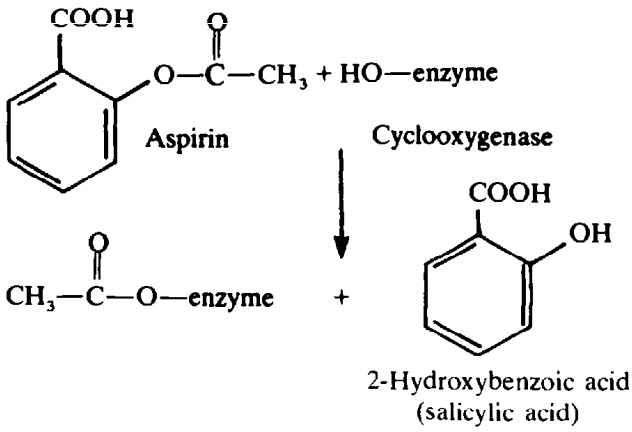
ganglioside cannot occur and the membrane lipid that is *turned over* in the lysosomes accumulates there (see Fig. 13-16).

PROSTAGLANDINS

13.12. What is the biochemical basis of the anti-inflammatory action of the drug aspirin?

SOLUTION

Prostaglandins such as PGE_2 are potent *vasodilators*, so they increase blood flow. Local vasodilation occurs at inflammatory sites. Aspirin irreversibly *inhibits* cyclooxygenase. A single serine residue in the enzyme is acetylated as follows:



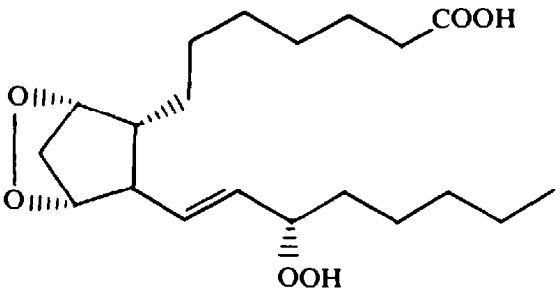
2-Hydroxybenzoic acid
(salicylic acid)

This inhibition of cyclooxygenase blocks the synthesis of prostaglandin, which in turn reduces the inflammatory response in the tissue.

13.13. Homo-γ-linolenic acid, like arachidonic acid, is a precursor of prostaglandin. How are the compounds that are derived from homo-γ-linolenic acid named?

$$CH_3—(CH_2)_4—CH=CH—CH_2—CH=CH—CH_2—CH=CH(CH_2)_6COOH$$

Homo-γ-linolenic acid



PGG_1

SOLUTION

Because homo-γ-linolenic acid lacks a double bond in position 5 but has a double bond in position 13, all derivatives of PGG_1, the primary prostaglandin of the homo-γ-linoleic acid series, are designated PG, followed by a letter A–I and the number 1 which denotes only one double bond in the molecule.

CHOLESTEROL AND RELATED COMPOUNDS

13.14. Cholecalciferol (vitamin D_3) is derived from cholesterol. How and in which tissue does the conversion occur?

SOLUTION

With normal exposure to sunlight, 7-dehydrocholesterol is converted to cholecalciferol in the skin.

13.15. Cholecalciferol is metabolized to produce steroid hormones. How and in which tissue does this occur?

SOLUTION

Cholecalciferol is hydroxylated at three positions in the carbon skeleton, 1, 24, and 25. In the liver, cholecalciferol is hydroxylated to 25-hydroxycholecalciferol. Further hydroxylation reactions occur in the kidney, resulting in the formation of three new metabolites. These are 1,25-dihydroxycholecalciferol; 24,25-dihydroxycholecalciferol; and 1,24,25-trihydroxycholecalciferol. 1,25-Dihydroxy- and 1,24,25-trihydroxycholecalciferol are active hormones involved in calcium uptake from the intestine.

REGULATION OF LIPID METABOLISM

13.16. How does the expenditure of ATP maintain the inactive form of liver HMG-CoA reductase? (Hint: The system of regulation is similar to that of glycogen phosphorylase, Section 11.9.)

SOLUTION

See Fig. 13-28. ATP is a cosubstrate in phosphorylation reactions that covalently modify enzymatic proteins. Two phosphorylation reactions are involved in the modifications of HMG-CoA reductase:

1. HMG-CoA reductase is phosphorylated in a reaction catalyzed by *HMG-CoA reductase kinase.*
2. *cAMP-dependent protein kinase* phosphorylates HMG-CoA reductase kinase, converting it to its active form. Subsequent dephosphorylation of HMG-CoA reductase and HMG-CoA reductase

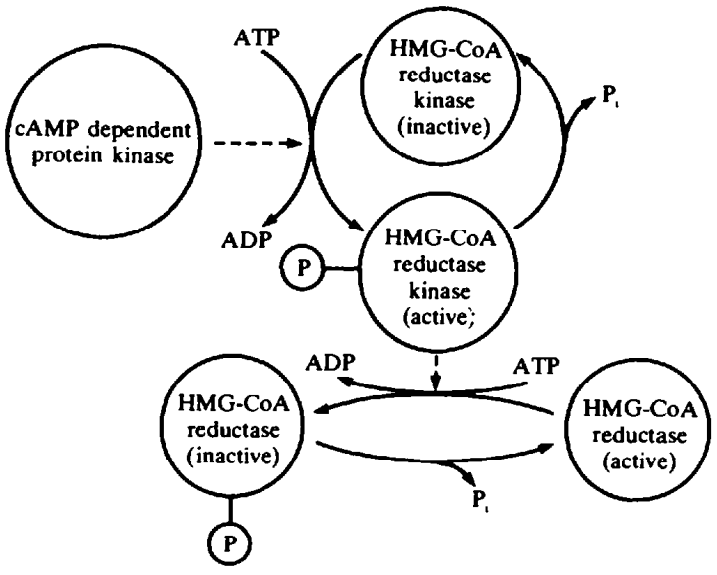


Fig. 13-28

kinase releases inorganic phosphate. Each HMG-CoA reductase molecule is phosphorylated at the expense of one molecule of ATP. Similarly, one molecule of ATP is utilized in the initial phosphorylation of HMG-CoA reductase kinase. However, one molecule of active HMG-CoA reductase kinase can catalyze the phosphorylation of many molecules of HMG-CoA reductase.

Supplementary Problems

13.17. What is the mechanism of the stimulation, by triglyceride in the diet, of the synthesis of cholesterol?

13.18. Why does feeding compounds that act as methyl-group donors to patients with fatty liver alleviate the condition?

13.19. How much energy, as ATP, is obtained from the complete oxidation of palmitic acid to carbon dioxide and water?

13.20. What is the yield of ATP from the oxidation of (*a*) 1 mole of acetoacetate and (*b*) 1 mole of 3-hydroxybutyrate?

13.21. What will be the contents of bile salt micelles after feeding: (*a*) 1,3-dipalmitoyl-2-linoleoylglycerol; (*b*) 1-stearoyl-2-arachidonylphosphatidylcholine?

13.22. How much glucose would be utilized to produce NADPH for the synthesis of one molecule of cholesterol from acetyl-CoA?

13.23. How is oleic acid converted to acetyl-CoA?

13.24. How does adipose tissue obtain the glycerol 3-phosphate necessary for triacylglycerol synthesis?

13.25. What properties of acetyl-CoA carboxylase are consistent with a regulatory role for the enzyme in fatty acid synthesis?

13.26. How do amino acids act as precursors in the synthesis of palmitic acid?

13.27. Which enzymes are involved in the oxidation of oleic acid to acetyl-CoA?

13.28. Explain how fat depots in mammalian adipose tissues are sources of intracellular water.

13.29. What effect would a deficiency of carnitine palmitoyltransferase in skeletal muscle have on the ability of a person to perform prolonged exercise?

13.30. Would aspirin be effective in preventing platelet aggregation?

13.31. If it were possible to provide a person with a diet free of cholesterol and triacylglycerol, how would this affect the deposition of triacylglycerol in adipose tissue?

13.32. 3-Hydroxy-3-methylglutaryl-CoA is an intermediate in the synthesis of acetoacetate and cholesterol in liver. How does this intermediate become available to each biosynthetic pathway?

13.33. How are cholesteryl esters synthesized?

13.34. Why does the concentration of ketone bodies in the blood increase during prolonged starvation?

13.35. What are the physical-chemical reasons that fatty acids are transported in lipoproteins?

13.36. In which cells may arachidonic acid be released into the cytosol?

Chapter 14

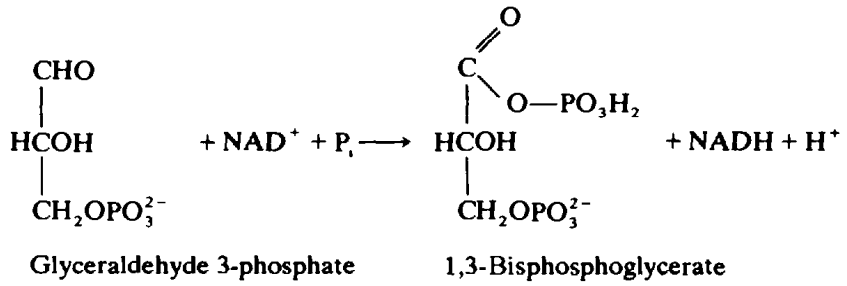
Oxidative Phosphorylation

14.1 INTRODUCTION

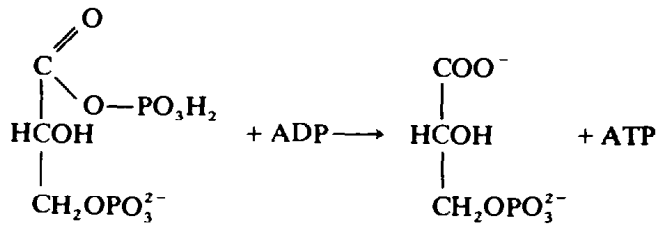
The study of *bioenergetics* involves the study of (1) the processes by which *reduced nicotinamide* and *flavin nucleotides*, generated primarily from the oxidation of carbohydrates (Chap. 11) and lipids (Chap. 13), are oxidized ultimately by molecular oxygen via the mitochondrial *electron-transport chain*, and (2) the mechanism by which this oxidation is coupled to *ATP synthesis*. The synthesis of ATP in this way is referred to as *oxidative phosphorylation*, in contrast to phosphorylation of ADP via soluble enzymes. The latter involves intermediate phosphate derivatives of the substrate and is known as *substrate-level phosphorylation* (Chap. 11).

EXAMPLE 14.1

In glycolysis, ADP is phosphorylated to ATP during the oxidation of glyceraldehyde 3-phosphate to 3-phosphoglycerate. The phosphorylated intermediate that receives the energy of the oxidation is 1,3-bisphosphoglycerate.



Glyceraldehyde 3-phosphate — 1,3-Bisphosphoglycerate



3-Phosphoglycerate

This is an example of *substrate-level phosphorylation*.

Oxidative phosphorylation is central to the metabolism of all higher organisms, because the *free energy* of hydrolysis of the ATP so generated is used in the synthesis of, inter alia, nucleic acids (Chaps. 7 and 16), proteins (Chaps. 4, 9, and 17), and complex lipids (Chap. 6), as well as in processes as diverse as muscle contraction (Chap. 5) and the transmission of nerve impulses.

14.2 COMPONENTS OF THE ELECTRON-TRANSPORT CHAIN

The *electron-transport chain*, or *respiratory chain* in mitochondria forms the means by which electrons, from the reduced electron carriers of intermediary metabolism, are channeled to oxygen and protons to yield H_2O. The main components of the chain are as follows.

$NAD^+/NADH$

The electron-transport reaction for the $NAD^+/NADH$ *conjugate redox pair* is:

$$NAD^+ + H^+ + 2e^- \rightarrow NADH \qquad E_0' = -0.32\,V$$

where E_0' is the standard redox potential (Chap. 10). In effect, electrons are transported as *hydride ions* (H^-), which are *formally* equivalent to ($H^+ + 2e^-$).

Question: How can electron transport by $NAD^+/NADH$ be measured?

$NADH$ has a characteristic light absorbance maximum at a wavelength of 340 nm, which is absent in NAD^+. Hence electron transport involving $NAD^+/NADH$ can be monitored by measuring the change in absorbance of a sample at 340 nm in a spectrophotometer.

EXAMPLE 14.2

NADH is oxidized by the electron-transport chain inside mitochondria. The inner mitochondrial membrane is not permeable to nucleotides. How can NADH generated in the cytoplasm (for example during glycolysis) participate in the electron-transport chain?

The reducing equivalents of cytosolic NADH are transferred into mitochondria via *shuttle mechanisms*, such as the one involving aspartate and malate, shown in Fig. 11-19 (page 333). The net effect of this shuttle is the transport of NADH *into* the mitochondrion.

Flavin Nucleotides

The electron-transport reactions for FAD and FMN (*f*lavin *mononu*cleotide) are:

$$FAD + 2H^+ + 2e^- \rightarrow FADH_2$$
$$FMN + 2H^+ + 2e^- \rightarrow FMNH_2$$

Electrons are effectively transported as H atoms by these nucleotides [$H \equiv (H^+ + e^-)$].

These carriers transfer electrons into the electron-transport chain independently of and bypassing the $NAD^+/NADH$ couple. The main shuttle for cytoplasmic reducing equivalents is the glycerol 3-phosphate shuttle that is shown in Fig. 11-20 (page 334).

EXAMPLE 14.3

Electrons from succinate, glycerol 3-phosphate, and the flavin-dependent steps of fatty acid oxidation enter the chain in the following way:

(*a*) Succinate + FAD → Fumarate + $FADH_2$

(*b*) Glycerol 3-phosphate + FAD → Dihydroxyacetone phosphate + $FADH_2$

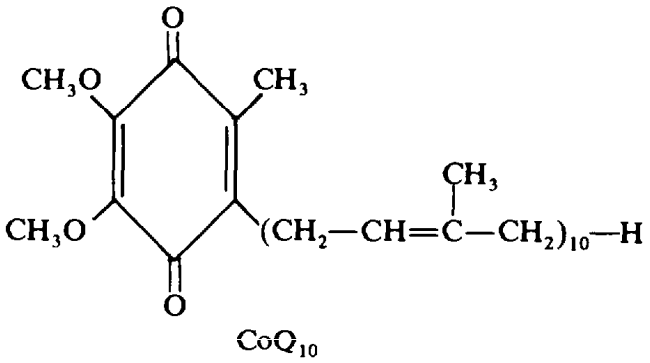
(*c*) Acyl-CoA + FAD → *trans*-Δ^2-Enoyl-CoA + $FADH_2$

Unlike the $NAD^+/NADH$ couple, the flavin coenzymes are not in free solution and are covalently linked to the dehydrogenases of their respective substrates. These enzymes are *membrane-bound* in close association with the respiratory chain in the inner mitochondrial membrane.

Coenzyme Q

Coenzyme Q (alternatively known as *ubiquinone* or *CoQ*) is a *benzoquinone* derivative with a long hydrocarbon side chain made up of repeating isoprene units. The number of units generally

categorizes the CoQ molecule. The form of CoQ in mammalian mitochondria contains 10 such units; hence, it is designated CoQ_{10}.



CoQ_{10}

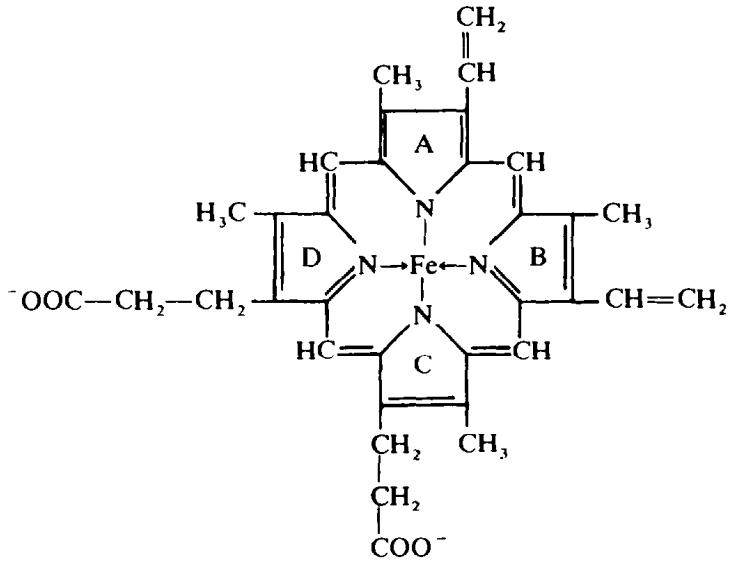
The molecule undergoes a $(2H^+ + 2e^-)$ reduction to form $CoQH_2$ (alternatively known as *ubiquinol*). The reduction may take place in two stages, yielding an intermediate, half-reduced, *free-radical* form (*semiquinone*), designated CoQH.

Cytochromes

The *cytochromes* (from the Greek: "cell colors") are a family of proteins containing prosthetic *heme groups* (see Chaps. 5 and 15). Mitochondria contain *three* classes of cytochromes: *a*, *b*, and *c*, which have hemes of different structures.

EXAMPLE 14.4

The general *cyclic tetrapyrrole* structure of a heme ring is shown below:



In cytochromes *c* and c_1, the heme ring is covalently attached to the protein via *thioether bonds*, formed by reaction of the *vinyl* groups ($-CH=CH_2$) on pyrrole rings A and B and cysteine residues of the protein. These thioether bonds are absent in cytochrome *b*. In cytochromes *a* and a_3, the vinyl group on ring A is replaced by a hydrocarbon chain, and the methyl group on ring D is replaced by a *formyl* ($-CHO$) group. In addition to heme, cytochromes *a* and a_3 contain bound Cu ions.

Question: How can cytochrome-mediated electron transport be measured?

Electron transport in cytochromes occurs by *direct electron transfer* between Fe^{2+} and Cu^{+} in cytochromes *a* and a_3. These changes in metal-ion oxidation state lead to changes in the visible absorption spectra of the cytochromes; spectrophotometric measurement of these changes allows quantification of the electron flow.

Iron-Sulfur Proteins

The electron-transport chain contains a number of *iron-sulfur proteins* (also known as *nonheme iron proteins*). The iron atoms are bound to the proteins via cysteine —S— groups and sulfide ions; one such 4-Fe cluster is shown in Fig. 14-1. These proteins mediate electron transport by direct electron transfer; changes in oxidation state of the iron in iron-sulfur proteins can be monitored by *electron spin resonance spectroscopy* (*ESR*).

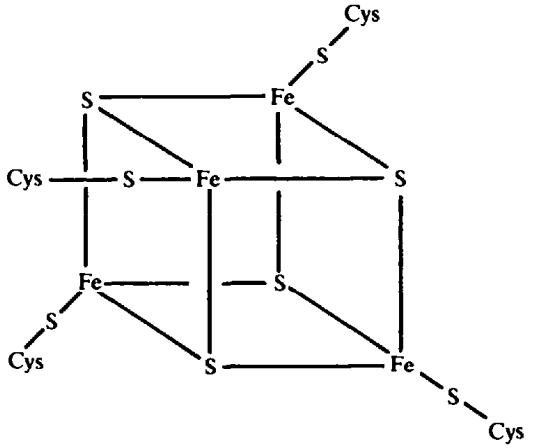


Fig. 14-1 Example of a 4-Fe iron-sulfur cluster.

14.3 ORGANIZATION OF THE ELECTRON-TRANSPORT CHAIN

The electron-transport chain is composed of the four complexes listed in Table 14.1. The pattern of electron transfer within these complexes is shown in Fig. 14-2.

Table 14.1. Protein Complexes of the Electron-Transport Chain

Complex	Enzymatic Function/Name	Functional Components
I	NADH/CoQ oxidoreductase	FMN; Fe-S clusters
II	Succinate/CoQ oxidoreductase (succinate dehydrogenase)	FAD; Fe-S clusters
III	CoQ-cytochrome *c* oxidoreductase	Cytochromes *b*, cytochrome c_1; Fe-S clusters
IV	Cytochrome *c* oxidase	Cytochromes *a* and a_3

Question: By what experimental means has this pattern of electron transfer been determined?

Two broad experimental approaches have been used: examination of the effects of *specific inhibitors*, which block electron flow through a particular complex; and use of synthetic *redox couples*, which are able to deliver electrons to specific complexes, depending on the relative E_0' values of the complex and the redox couple.

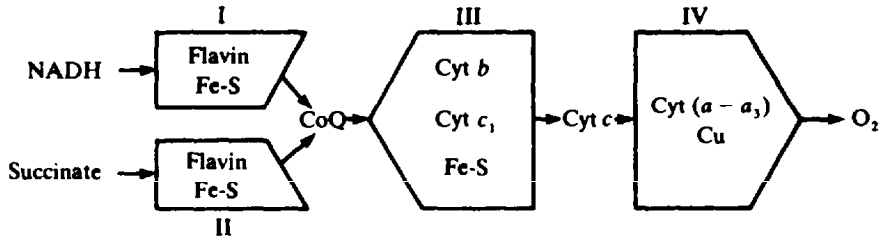


Fig. 14-2 Schematic organization of the electron–transport chain in mitochondria.

EXAMPLE 14.5

The sites of action of some of the commonly used inhibitors of the electron-transport chain are shown in Fig. 14-3. These sites have been established by application of the *crossover theorem* (Chap. 10). For example, the fungus-derived antibiotic *antimycin A* causes an *increase* in the level of reduced cytochrome *b* and a *decrease* in the level of reduced cytochrome c_1 (i.e., an increase in the level of oxidized cytochrome c_1); thus, it is inferred that antimycin A interacts with complex III.

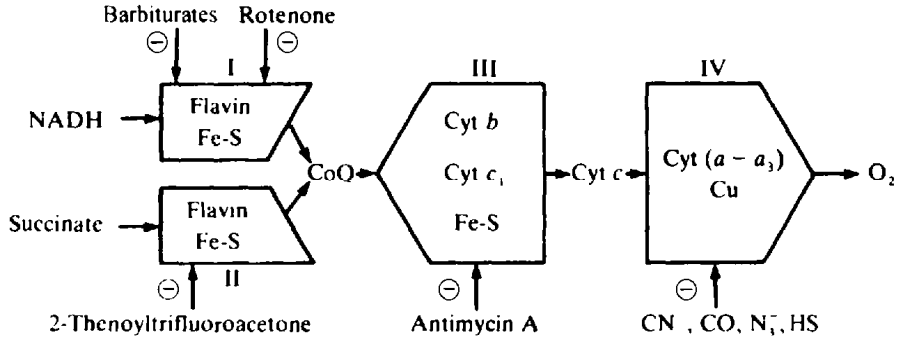


Fig. 14-3 Sites of action of inhibitors of the electron-transport chain.

Question: How can the E_0' values of electron-transport-chain components be measured?

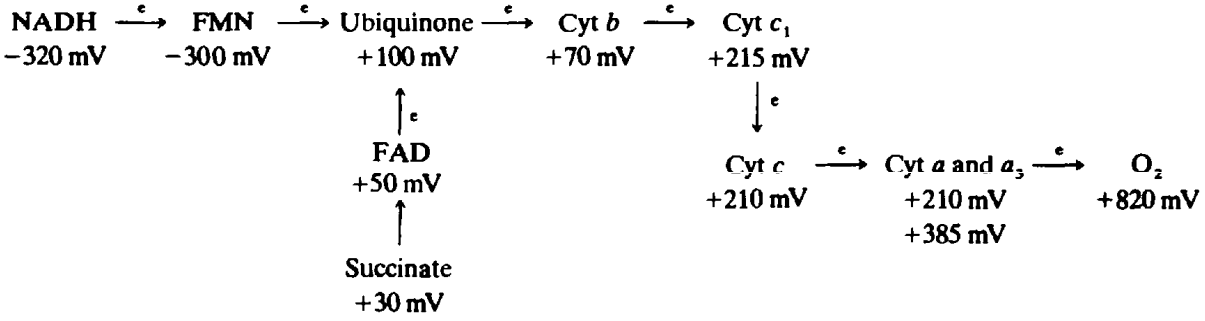
Measurements can be done using the technique of *redox potentiometry*. In experiments of this type, mitochondria are incubated anaerobically in the presence of a reference electrode [for example, a hydrogen electrode (Chap. 10)] and a platinum electrode and with *secondary redox mediators*. These mediators form redox pairs with E_0' values intermediate between the reference electrode and the electron-transport-chain component of interest; they permit rapid equilibration of electrons between the electrode and the electron-transport-chain component. The experimental system is allowed to reach equilibrium at a particular E value. This value can then be changed by addition of a reducing agent (such as reduced ascorbate or NADH), and the relationship between E and the levels of oxidized and reduced electron-transport-chain components is measured. The E_0' values can then be calculated using the Nernst equation (Chap. 10):

$$E = E_0' - \frac{RT}{nF} \ln \frac{[A_{red}]}{[A_{ox}]} \qquad (14.1)$$

where $[A_{red}]$ and $[A_{ox}]$ are the concentrations of reduced and oxidized electron-chain component A; these are obtained from spectroscopic measurements of the mitochondrial preparation.

EXAMPLE 14.6

Studied by redox potentiometry, the components of the electron-transport chain have been assigned the E_0' values shown below:



This technique also indicates (1) whether a component is a $1e^-$ carrier or a $2e^-$ carrier, which is evident from the slope of a plot of E versus $\ln\dfrac{[A_{red}]}{[A_{ox}]}$; and (2) whether a component is an $(H^+ + e^-)$ carrier, in which case the value of E_0' is pH-dependent.

14.4 COUPLING OF ELECTRON TRANSPORT AND ATP SYNTHESIS

It is now generally accepted that the coupling of electron transport and ATP synthesis is brought about by the action of a *proton electrochemical-potential gradient*, denoted by the symbol $\Delta\mu_{H^+}$. This gradient arises as a consequence of electron transport and is dissipated by *ATP synthase* to generate ATP from ADP and P_i.

Question: How is $\Delta\mu_{H^+}$ defined experimentally?

It is defined as

$$\Delta\mu_{H^+} = \Delta\psi - \frac{2.3RT}{F}\Delta\text{pH} \qquad (14.2)$$

where $\Delta\psi$ and ΔpH are, respectively, the electrical potential (in volts) and pH difference across the inner mitochondrial membrane; R, T, and F are the gas constant, absolute temperature, and Faraday constant, respectively. Both $\Delta\psi$ and ΔpH can be measured experimentally.

EXAMPLE 14.7

The most common way in which $\Delta\psi$ is determined is from measurements of the concentrations inside and outside mitochondria, at equilibrium, of an ionizable compound that is permeable to the inner mitochondrial membrane. $\Delta\psi$ can then be calculated using the Nernst equation (*14.1*), written in the form:

$$\Delta\psi = \frac{RT}{nF}\ln\frac{[X^{n+}]_{out}}{[X^{n+}]_{in}} \qquad (14.3)$$

for an ion X of charge $+n$. It can also be calculated from changes in the spectral properties of membrane constituents arising from $\Delta\psi$ (*electrochromism*); a value of 200 mV for $\Delta\psi$ corresponds to an electric field across the membrane of $\sim 3 \times 10^5\,\text{V cm}^{-1}$, given a membrane thickness of 6 nm.

EXAMPLE 14.8

The value of ΔpH across the inner mitochondrial membrane can be estimated from the equilibrium distribution of *electroneutrally-permeant weak acids* (or *weak bases*). The logic underlying such experiments is illustrated below.

For the reaction

$$H^+_{out} + A^-_{out} \overset{K_a}{\rightleftharpoons} HA_{out} \rightleftharpoons HA_{in} \overset{K_a}{\rightleftharpoons} H^+_{in} + A^-_{in}$$

the neutral species (HA) will equilibrate across the membrane independently of $\Delta\psi$. Thus, at equilibrium

$$[HA]_{in} = [HA]_{out} \tag{14.4}$$

Hence, assuming that the equilibrium constant K_a is the same for the ionization of HA inside and outside the mitochondrion, and the value of K_a is sufficiently low that $[HA] \approx 0$ and it will not contribute to the measurements, then

$$\frac{[H^+]_{out}[A^-]_{out}}{[HA]} = \frac{[H^+]_{in}[A^-]_{in}}{[HA]} \tag{14.5}$$

and

$$\frac{[A^-]_{out}}{[A^-]_{in}} = \frac{[H^+]_{in}}{[H^+]_{out}} \tag{14.6}$$

Therefore, ΔpH can be calculated from measurements of $[A^-]_{out}$ and $[A^-]_{in}$.

14.5 THE RATIO OF PROTONS EXTRUDED FROM THE MITOCHONDRION TO ELECTRONS TRANSFERRED TO OXYGEN

The *chemiosmotic model* requires that flow of electrons through the electron-transport chain leads to *extrusion* of protons from the mitochondrion, thus generating the proton electrochemical-potential gradient. Measurements of the number of H^+ ions extruded per O atom reduced by complex IV of the electron-transport chain (the H^+/O ratio) are experimentally important because the ratio can be used to test the validity of *mechanistic models* of proton translocation (Sec. 14.6).

EXAMPLE 14.9

The H^+/O ratio can be measured by incubating mitochondria with an appropriate substrate (for example, NADH or succinate) under *anaerobic* conditions. The reaction is initiated by addition of a known amount of O_2, and H^+ extrusion is measured (as ΔpH) by using a pH electrode. The results of such an experiment are shown in Fig. 14-4.

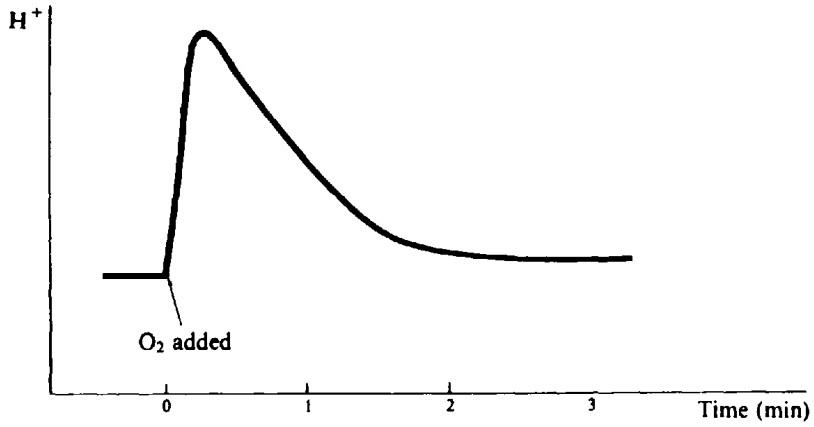


Fig. 14-4 Proton extrusion during respiration.

There are a number of complications in the interpretation of such data. The mitochondria must be permeable to cations other than H^+. If this were not the case, then H^+ extrusion would lead to an increase in $\Delta\psi$, which would diminish further H^+ extrusion. Incorporation of K^+ and addition of the *ionophore* valinomycin to the mitochondrial preparation prevents increases in $\Delta\psi$.

Other complications in analyzing the data from this experiment include the necessity to correct for movement of H^+ back into the mitochondria; this may occur *directly* or via the *Na^+/H^+antiport* translocator of the mitochondrial inner membrane, or via the *H^+/phosphate symport* translocator (see Prob. 14.7).

Question: How does valinomycin render mitochondrial membranes permeable to K^+?

Valinomycin, an antibiotic, is a *mobile* K^+ carrier with an interior rich in polar amino acid residues (providing a binding site for K^+) and hydrophobic valine residues on the outside. The latter allow valinomycin, bearing K^+ in its interior, to diffuse across membranes, thus rendering them permeable to K^+. Free K^+ is unable to diffuse at appreciable rates across membranes.

14.6 MECHANISTIC MODELS OF PROTON TRANSLOCATION

Loop Mechanisms

In the chemiosmotic model, as first developed by Mitchell in the early 1960's, proton translocation arises from transfer of electrons from an $(H^+ + e^-)$ carrier (such as $FMNH_2$) to an electron carrier (such as an iron-sulfur protein), with expulsion of protons to the outer compartment of the inner mitochondrial membrane. This process is followed by electron transfer to an $(H^+ + e^-)$ carrier, with uptake of protons from the matrix. In this model, the electron-transport chain is organized into *three* such loops, as shown in Fig. 14-5.

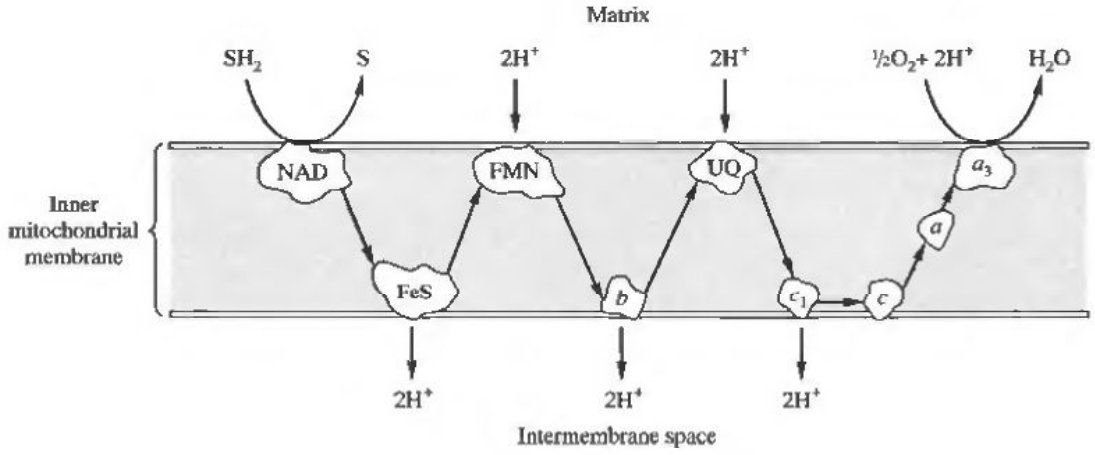


Fig. 14-5 Mitchell loop mechanism for proton translocation. SH_2 = reduced substrate, S = oxidized substrate; FeS = iron-sulfur protein; a, a_3, b, c, c_1 = cytochromes; UQ = coenzyme Q (ubiquinone).

EXAMPLE 14.10

What experimentally testable predictions does the Mitchell loop mechanism make?

This scheme makes two experimentally testable predictions:

1. There should be a ratio of 2 H^+ ions translocated per 2 electrons transported in each loop.
2. The electron-transport chain should be organized to give alternating $(H^+ + e^-)$ and pure e^- carriers.

The experimental evidence pertaining to prediction (1) is discussed later in this section, while problems relating to (2) are outlined in Example 14.11.

Proton-Pump Mechanisms

In these mechanisms, electron transport through the various components of the electron-transport chain leads to structural changes in the proteins of the chain, such that changes in their pK_a values (Chap. 3) of ionizable amino acid residues occurs. For example, an *increase* in the pK_a of a residue adjacent to the *matrix* side of the membrane would lead to proton *uptake* from the matrix, while a *decrease* in the pK_a of a residue adjacent to the *intermembranous* side of the membrane could lead to *release* of a proton. The net effect of these processes is the transfer of protons from the matrix to the intermembranous side of the membrane. However, proton-pump mechanisms do *not* make strong predictions of the H^+/e^- stoichiometries.

Proton Translocation by Complex I

Complex I mediates electron transfer from $NADH + H^+$ to coenzyme Q. The consequent reduction of coenzyme Q is coupled to *vectorial* proton transport; the H^+/e^- stoichiometry (ratio) has been variously estimated to be 1.0, 1.5, and 2.0.

Question: Is an H^+/e^- stoichiometry of 2 consistent with the simple chemiosmotic hypothesis?

No, it is inconsistent with this hypothesis, which predicts an H^+/e^- value of 1.0. Such stoichiometries may, however, be explained by *proton-pump mechanisms*, in which electron transfer is coupled to changes in the pK_a values of proteins within Complex I.

Proton Translocation by Complex II

Complex II mediates electron transfer from succinate to coenzyme Q. This process does *not* appear to be coupled to vectorial proton translocation.

Proton Translocation by Complex III

Complex III catalyzes electron transfer from reduced coenzyme Q to cytochrome *c*; this process is coupled to vectorial proton translocation with an H^+/e^- stoichiometry of 2.

EXAMPLE 14.11

How can the above H^+/e^- stoichiometry be explained?

This has been one of the most controversial areas of bioenergetics and is concerned with the role of coenzyme Q. The simplest view of the role of this coenzyme is that it acts as a *mobile* ($2H^+ + 2e^-$) *carrier*, linking complexes I and II with complex III. However, coenzyme Q may be involved in ($H^+ + e^-$) transfer within complex III. One model for this is the *proton-motive Q cycle* (Fig. 14-6), developed by Mitchell in 1975. This model satisfies prediction (2) of Example 14.10, in that coenzyme Q acts as an ($H^+ + e^-$) carrier in *two* loops. In this model, *reduced* coenzyme Q (QH_2) is linked to *oxidized* coenzyme Q (Q) via the free-radical *semiquinone* ($QH\cdot$). This model provides an explanation for the H^+/e^- stoichiometry.

Proton Translocation by Complex IV

Complex IV catalyzes electron transfer from cytochrome *c* to O_2; this process appears to be coupled to proton translocation, with an H^+/e^- value of 2. Two models have been developed to account for these values (Fig. 14-7). Current understanding is that complex IV is capable of acting as a proton pump.

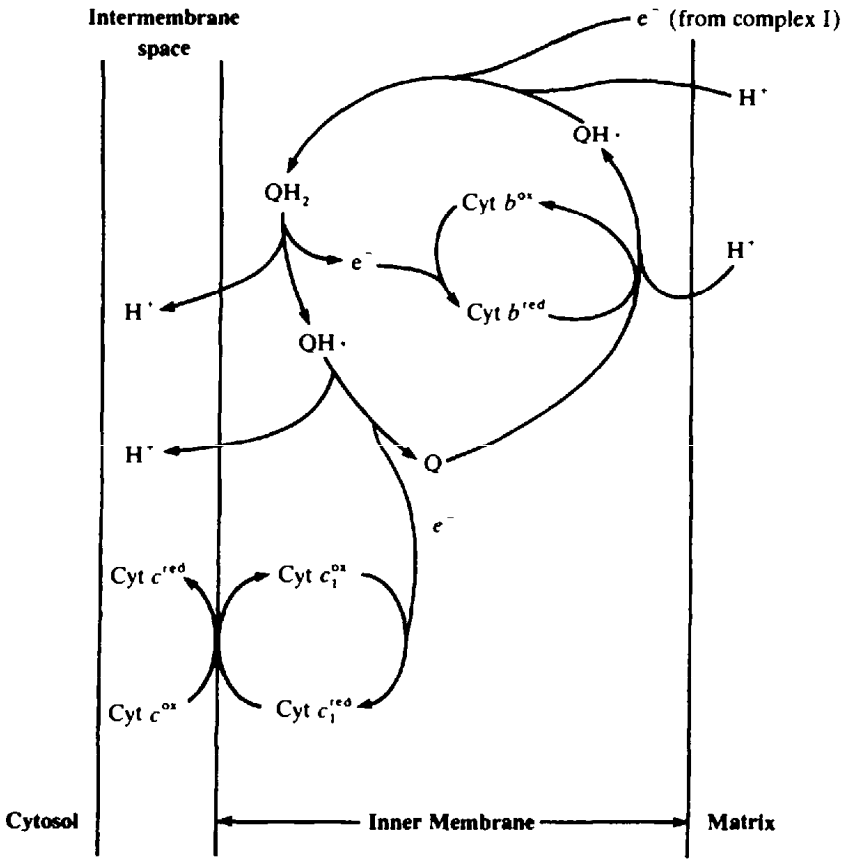


Fig. 14-6 Representation of the proton-motive Q cycle.

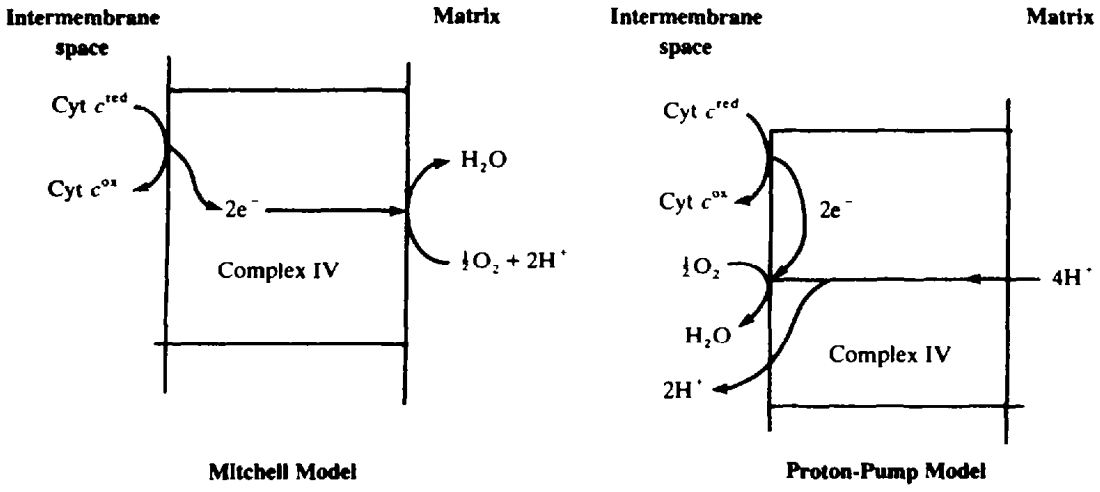


Fig. 14-7 Two models for proton translocation.

Question: How could the Mitchell model of oxidative phosphorylation give an H^+/e^- value of 1 in a situation when no protons traversed the membrane?

The Mitchell model can be described by the equation

$$2e^- + 2H^+ + \tfrac{1}{2}O_2 \rightarrow H_2O$$

Thus, there would be a net decrease in $[H^+]$ on the matrix side of the mitochondrial inner membrane simply from this chemical reaction.

14.7 ATP SYNTHASE

ATP synthase is found in *all* energy-transducing membranes, including those of *mitochondria*, *chloroplasts*, and *bacteria*. The enzyme complex most probably evolved in primitive anaerobic life forms, before the appearance of oxygen-based respiration. In those early organisms it possibly functioned in reverse, utilizing ATP generated by *fermentation* reactions to create proton gradients that were coupled to ion-transport systems.

Electron micrographs of mitochondria show globular structures protruding from the matrix side of the inner mitochondrial membrane. These globular units can be detached from the membranes by relatively mild treatment such as low ionic strength or urea. The solubilized particles, known as the F_1 component of ATP synthase, can hydrolyze ATP, but cannot synthesize it, and hence has been termed an *ATPase*.

EXAMPLE 14.12

The F_1 component of ATP synthase comprises 5 separate polypeptide chains, designated α ($M_r = 56{,}000$), β ($M_r = 53{,}000$), γ ($M_r = 33{,}000$), δ ($M_r = 14{,}000$), and ε ($M_r = 6{,}000$), in the stoichiometry $\alpha_3\beta_3\gamma\delta\varepsilon$. The catalytic site for ATPase activity appears to reside on the β chains, while the δ and ε chains appear to be involved in attaching the F_1 complex to the membrane.

ATP synthase activity can be restored by adding back the F_1 complex to the depleted membranes. The F_1 complexes bind to membrane channels known as the F_0 complex, which are also composed of multiple subunits. The polypeptides of the F_0 component are very hydrophobic and form a proton transport channel through the membrane, which links the proton gradient to ATP synthesis. This channel appears to be lined with hydrophilic residues such as seryl, threonyl and carboxyl groups. The "stalk" that connects the F_1 to the F_0 complex comprises one copy each of the polypeptide known as the *oligomycin-sensitivity-conferring protein* (OSCP) and another protein known as F_6.

EXAMPLE 14.13

The major subunit of the F_0 complex is a small ($M_r = 5{,}400$) polypeptide referred to as *proteolipid* because of the high proportion of phospholipid bound to it. This proteolipid, now sequenced from a number of sources, forms a *hairpin loop* that traverses the inner mitochondrial membrane and has unique hydrophobic amino acid sequences flanking a central, short, highly charged segment that interacts with the OSCP or with F_1 component δ.

The compound *dicyclohexylcarbodiimide* (DCCD) inhibits proton translocation through F_0 by reacting with the carboxyl group of a single glutamate residue in the channel-forming loop of the proteolipid.

Oligomycin is an antibiotic that inhibits respiration in intact mitochondria. Respiration is not inhibited in *uncoupled* mitochondria, i.e., those mitochondria in which O_2 consumption occurs but in which no ATP is synthesized. Thus, oligomycin does not block respiratory carriers, in contrast to inhibitors such as rotenone and cyanide. Instead, oligomycin blocks proton translocation through the F_0 component to the F_1 component, through a specific interaction with a subunit of the membrane-associated F_0. The subscript "o" in the term F_0 was originally used to indicate the "oligomycin-sensitive" complex.

14.8 THE MECHANISM OF ATP SYNTHESIS

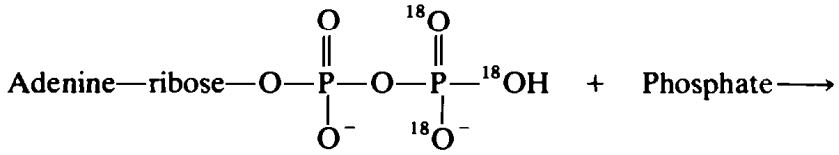
The main tenet of the chemiosmotic hypothesis is that a transmembrane electrochemical gradient acts as the *intermediate* in the transfer of energy to ATP; this energy is available from the difference in redox potential between the $NAD^+/NADH$ couple and the $O_2/2H_2O$ couple in the respiratory chain. Energy coupling occurs in *each* of the three respiratory complexes I, III, and IV, where in each case there is sufficient difference in *midpoint potential* of the donor and acceptor carriers to drive proton transport *against* the electrochemical gradient. Thus, oxidation of NADH results in the production of approximately three ATP molecules per atom of O reduced to water (a P/O ratio of 3). Oxidation of succinate, on the other hand, yields only two ATP molecules, a P/O ratio of 2.

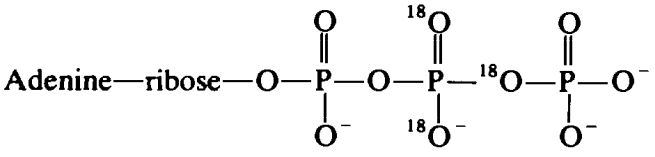
The electrochemical gradient is the driving force in the reaction:

$$ADP + P_i + \Delta\mu_{H^+} \rightleftharpoons ATP + H_2O$$

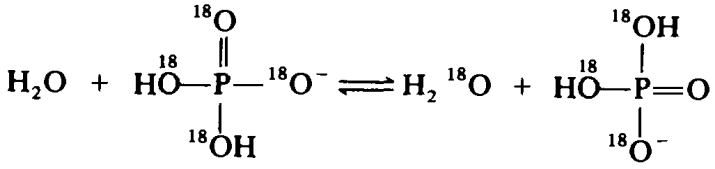
The x-ray crystal structure of the F_1 ATPase has recently been determined at 2.8 Å resolution. The three α and three β chains are arranged alternately in a circle around an α-helical coiled coil of the γ subunit that protrudes into the stalk region. The F_1 subunit is fundamentally asymmetrical because there is only a single γ subunit to three α and three β subunits; in addition, each of the three α chains and each of the three β chains adopts a slightly different conformation. One β subunit binds ATP, one binds ADP, while the third appears to have an empty nucleotide binding site.

The concept of a high-energy phosphorylated intermediate at "coupling sites" was a feature of all earlier concepts of the mechanism of ADP phosphorylation. The *possibility* of such a precursor, engendered by the electrochemical gradient on F_1, has however been eliminated by studies involving the use of ^{18}O-labeled components of the reaction:





A possibility of a high-energy derivative of an F_1 component involving an *anhydride of ADP* has been excluded by the demonstration that the oxygen bridge joining the β and γ phosphates of ATP is derived from the ADP-OH and not the P_i-OH. A high-energy derivative of an F_1 component involving an *anhydride of phosphate* has also been excluded as a possibility by the demonstration that exchange of ^{18}O between P_i and water, as indicated below, is enhanced and not reduced by addition of ADP.



Covalent bond formation during ATP synthesis is thus restricted to the one linking the β and γ phosphates of ATP itself.

The energy of the electrochemical $\Delta\mu_{H^+}$ appears not to be utilized *directly* in the making of this covalent linkage but is used, rather, in the binding of ADP and P_i to F_1 and in the subsequent release of ATP.

EXAMPLE 14.14

Models based on the above concept have been proposed, such as the one shown in Fig. 14-8. Evidence for this model is based on studies of the hydrolysis of ATP labeled in the γ phosphate with ^{18}O. The label is lost from the released P_i to water ($H^{18}OH$) because of the rapid reversibility of reactions 2 and -2. Uncoupling agents do *not* influence this loss, indicating μ_{H^+} is not involved in reactions 2 or -2, nor in reactions -3 or -1.

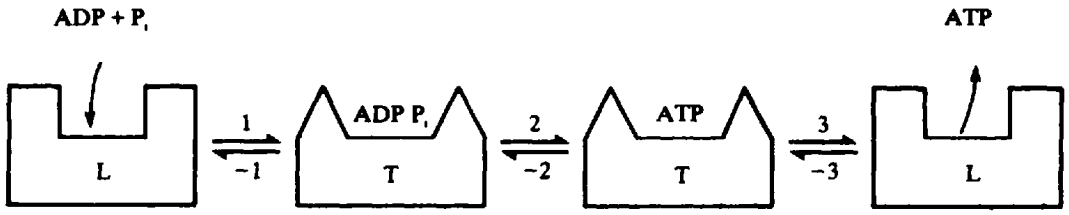


Fig. 14-8 A model for binding of ADP and P_i, with subsequent release of ATP. L = loose binding complex; T = tight binding complex.

Energy is thus required in the *binding* of ADP and P_i, which converts the loose (L) binding complex to a tight (T) binding complex. Reversible loss of water (ATP synthesis) or gain of water (ATP hydrolysis) in the tight binding complex may occur without involvement of the μ_{H^+}. Energy is also required in the conversion of the T-bound ATP complex to an L-bound ATP complex, from which ATP is released without additional energy involvement.

On the basis of the above arguments and the recent x-ray crystal structure of the ATPase, the most likely mechanism involves cyclic conformational changes, and perhaps rotation of subunits, driven by the transport of protons through the F_o complex (Fig. 14-9). The three β subunits appear to exist in three separate states: a *tight* state, T, which binds ATP tightly, and in which ADP is phosphorylated, a *loose* state, L, to which ADP and inorganic phosphate can bind, and an *open* state, O, from which ATP can be released. Once ADP and P_i are bound to the L state, a conformation change brought about by proton translocation switches the L to the T state, allowing the next ATP to be formed, simultaneously switching the adjacent T subunit to the O state, and allowing the previously formed ATP to be released. The third β chain switches from the O to the L state, to allow ADP to bind for the next round of ATP synthesis.

Tight coupling between ATP synthesis and proton translocation is dependent on the impermeability of the membrane to protons, so that the F_o channel and ATP synthesis provide the only way for protons to reenter the mitochondrial matrix. Physical damage to the membranes, or chemicals that allow the dissipation of the proton or electrical potential gradient, will allow alternative pathways for reentry of protons, and will *uncouple* respiration from ATP synthesis (see Problem 14.5).

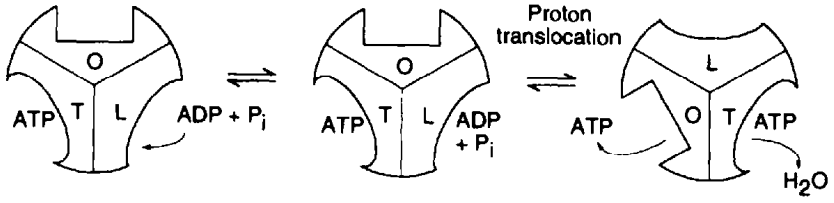


Fig. 14-9 A model for the cyclic synthesis of ATP by ATP synthase coupled to conformation changes in the β subunits brought about by proton translocation through the F_o complex. ATP synthesis occurs in the "tightly bound" (T) state, but the ATP can only be released from the "open" (O) state. The energy of the electrochemical gradient is used to switch the T to the O state. The third state, L, can bind ADP.

14.9 TRANSPORT OF ADENINE NUCLEOTIDES TO AND FROM MITOCHONDRIA

Although the mechanism of ATP synthesis is not fully understood, there is no doubt that the ATP is synthesized in the matrix of the mitochondrion. Mitochondrial ATP is then *exported* to the cytoplasm. A specific carrier (*translocase*; M_r = 40,000) is involved in the simultaneous transport of

ATP *out* of and ADP *into* the mitochondrion. Translocation is inhibited by two well-known toxins, *atractyloside* and *bongkrekic acid*, the former a glucoside found in the Mediterranean thistle and the latter produced by a *Pseudomonas* bacterium.

EXAMPLE 14.15

The translocator moves the two nucleotides in either direction. However, ATP is transported as ATP^{4-} and ADP as ADP^{3-}. Thus the equilibrium position of the exchange is dependent on the electrochemical potential difference across the membrane. When the electrical potential difference is 160 mV, the ratio of ATP/ADP in the medium outside the mitochondrion is 125/1.

Solved Problems

COMPONENTS OF THE ELECTRON-TRANSPORT CHAIN

14.1. A suspension of mitochondria was incubated under anaerobic conditions in the presence of NADH. Oxidative phosphorylation was initiated by a pulse of oxygen, and it was found that the ATP concentration in the suspension increased by 1.5×10^{-4} mol L^{-1}, with a corresponding decrease in the absorbance of the suspension of 0.30 at 340 nm, when using a 1-cm path length in a spectrophotometer. Given that the absorbance coefficient of NADH at 340 nm is 6.2×10^3 L mol^{-1} cm^{-1}, calculate the number of ATP molecules produced per molecule of NADH oxidized.

SOLUTION

The Beer-Lambert equation (Chap. 3) is

$$A = \varepsilon C l$$

where A is the absorbance of the solution; ε, the absorbance coefficient; C, the concentration in mol L^{-1}; and l, the length of light path in cm. By rearranging the equation and substituting the values given above, we obtain

$$\Delta C = \frac{\Delta A}{\varepsilon l}$$

$$\Delta[\text{NADH}] = \frac{\Delta A_{340\text{nm}}}{\varepsilon_{340\text{nm}} \times 1\text{ cm}}$$

$$= \frac{0.30}{6.2 \times 10^3\text{ L mol}^{-1}\text{ cm}^{-1} \times 1\text{ cm}} = 4.84 \times 10^{-5}\text{ mol L}^{-1}$$

Comparison of this value with the change in [ATP] (1.5×10^{-4} mol $L^{-1} \div 4.84 \times 10^{-5}$ mol L^{-1}) gives a value of 3.1 ATP produced per NADH oxidized; this is close to the theoretically predicted value of 3.0.

14.2. Muscle tissue contains two glycerol-3-phosphate dehydrogenases: a cytosolic enzyme, which uses NADH, and a flavin-nucleotide-dependent mitochondrial enzyme. What is the metabolic significance of these two enzymes?

SOLUTION

These two enzymes provide a *shuttle mechanism* for transport of reducing equivalents of NADH (generated during glycolysis in the cytosol) into the mitochondria. The cytosolic enzyme catalyzes the following reaction:

$$\text{Dihydroxyacetone phosphate} + \text{NADH} + \text{H}^+ \rightleftharpoons \text{NAD}^+ + \text{Glycerol 3-phosphate}$$

The glycerol 3-phosphate traverses the mitochondrial membrane and is then oxidized back to dihydroxy-acetone phosphate by the mitochondrial enzyme:

$$\text{Glycerol 3-phosphate} + \text{FAD} \rightleftharpoons \text{Dihydroxyacetone phosphate} + \text{FADH}_2$$

The $FADH_2$ enters the electron-transport chain at coenzyme Q, while the dihydroxyacetone phosphate can return to the cytoplasm. Although this shuttle is generally *inefficient*, in the sense that only two ATP molecules are produced per $FADH_2$ molecule oxidized, compared with three for NADH oxidation, it provides a mechanism for regeneration of NAD^+ in the cytosol. The presence of cytosolic NAD^+ is essential for continued glycolysis (see Fig. 11-20).

14.3. How can the two types of sulfur atoms in an iron-sulfur protein be distinguished experimentally?

SOLUTION

The sulfur atoms linking the Fe atoms can be removed from the iron-sulfur protein at low pH; they are *acid-labile*, while the sulfur atoms derived from cysteine residues are covalently attached to the β carbon of the cysteine residues and are *not* acid labile.

14.4. In the succinate dehydrogenase-catalyzed reaction, why is the appropriate electron acceptor FAD rather than NAD^+, which is used in the other redox reactions of the citric acid cycle (Chap. 12)?

SOLUTION

For FAD-mediated oxidation of succinate, the half-reactions are:

$$\text{FAD} + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{FADH}_2 \qquad E_0' = 0.05\ \text{V}$$

$$\text{Fumarate} + 2\text{H}^+ + 2\text{e}^- \rightarrow \text{Succinate} \qquad E_0' = +0.03\ \text{V}$$

Hence, for the overall reaction

$$\text{Succinate} + \text{FAD} \rightarrow \text{Fumarate} + \text{FADH}_2$$

E_0' is 0.02 V. The value of $\Delta G^{\circ\prime}$ can be calculated as follows:

$$\Delta G^{\circ\prime} = -nFE_0'$$
$$= -2 \times 96.5 \times 0.02 = -3.86\ \text{kJ mol}^{-1}$$

For NAD^+-mediated oxidation of succinate, the nucleotide E_0' value is -0.32 V, giving an E_0' value for the whole reaction of -0.35 V. Hence the $\Delta G^{\circ\prime}$ for NAD^+-mediated oxidation of succinate is unfavorable, being $+67.6\ \text{kJ mol}^{-1}$. Recalling that

$$\Delta G^{\circ\prime} = -RT \ln K$$

the equilibrium constant for FAD-mediated oxidation of succinate is $\sim 10^{12}$-fold more favorable than that for the NAD^+-mediated oxidation.

COUPLING OF ELECTRON TRANSPORT AND ATP SYNTHESIS

14.5. *Uncoupling agents* are compounds that prevent ATP synthesis in mitochondria but allow electron transport to proceed. They generally act by increasing the permeability of the inner mitochondrial membrane to H^+, thus dissipating the H^+ gradient. A widely used uncoupling agent is 2,4-dinitrophenol. How could this compound increase the permeability of the inner mitochondrial membrane to H^+?

SOLUTION

At physiological pH, 2,4-dinitrophenol exists predominately as the anion, $C_6H_4(NO_2)_2O^-$. The membrane is permeable both to this anion and to the protonated form, $C_6H_4(NO)_2OH$. The latter form can carry protons across the membrane and return in the anionic form to be reloaded with a proton. Thus, 2,4-dinitrophenol can dissipate the H^+ gradient.

14.6. Addition of DCCD (dicyclohexylcarbodiimide) to mitochondrial preparations *decreases* the rates of both ATP synthesis and electron transport. Only the latter process can be restored to normal levels upon addition of 2,4-dinitrophenol. How can these observations be explained?

SOLUTION

DCCD inhibits proton translocation through the F_0 subunit of the ATP synthase. Thus, the value of $\Delta\mu_{H^+}$ increases to a point where proton translocation, and hence electron transport, becomes thermodynamically unfavorable. In addition, DCCD inactivates the ATP synthesis function of the ATP synthase. The uncoupler, 2,4-dinitrophenol, renders the inner mitochondrial membrane permeable to protons, leading to a decrease in the value of $\Delta\mu_{H^+}$ and restoration of electron transport. However, 2,4-dinitrophenol cannot restore the activity of the DCCD-treated ATP synthase.

14.7. Most mitochondria contain an active H^+/phosphate symport protein containing a sulfhydryl group essential for its translocator activity. (*a*) How could this translocator complicate measurements of the H^+/O ratio (as in Example 14.9), and (*b*) how could the presence of the sulfhydryl group be used to obviate these complications?

SOLUTION

(*a*) The protons extruded from the mitochondria may be transported, together with phosphate, back into the mitochondria via the H^+/phosphate translocator. This would lead to an underestimate of the number of protons extruded and, hence, an underestimation of the H^+/O ratio.

(*b*) Because the H^+/phosphate translocator has an essential sulfhydryl group, treatment of mitochondrial preparations with *group-selective reagents*, such as *N-ethylmaleimide*, that react with sulfhydryl groups (Chap. 3) will lead to selective inactivation of the H^+/phosphate translocator. Hence, reentry of protons mediated by this translocator will not occur. Experimentally, higher H^+/O ratios are obtained from mitochondria treated with *N*-ethylmaleimide than from untreated mitochondria.

14.8. *Brown fat* is a form of adipose tissue found under the skin on the backs of many young animals. Mitochondria from this tissue have a P/O ratio of less than 1 for ATP synthesis arising from oxidation of NADH. What may be the physiological function of brown fat tissue?

SOLUTION

From the above P/O ratio, it is clear that ATP synthesis in brown fat mitochondria is *naturally uncoupled* from electron transport. Hence, protons extruded from the mitochondria during electron transport must reenter without concomitant ATP synthesis. The energy released as heat during this reentry may help to keep the young animals warm. Such small organisms have a high *surface-to-volume ratio* and therefore readily lose heat through convective and radiative processes.

14.9. In an experiment, purified F_0 protein can be incorporated into synthetic phospholipid vesicles. When these vesicles are preloaded with K^+ and then valinomycin added to the suspension, they are able to *take up* protons. No H^+ uptake was seen with control phospholipids vesicles containing no F_0 protein. What might be the basis of these observations?

SOLUTION

The addition of *valinomycin*, a K^+ ionophore, leads to generation of a *diffusion potential* through efflux of K^+ down its concentration gradient. The F_o protein permits uptake of H^+ by the vesicles (which are impermeable to H^+ in the absence of F_o) in response to this diffusion-generated gradient.

14.10. In an experiment, a suspension of mitochondria provided with adequate oxygen and pyruvate, but no ADP, consumed oxygen at a very low rate. The relative states of reduction of components of the electron-transport chain were determined: NAD, 100 percent; coenzyme Q, 40 percent; cytochrome *b*, 38 percent; cytochrome *c*, 14 percent; and cytochrome *a*, 0 percent. How could these data allow definition of the sites of oxidative phosphorylation?

SOLUTION

The *absence* of ADP is acting, in effect, as an *inhibitor* of electron transport, for reasons discussed in Prob. 14.6 below. Hence, by application of the *crossover theorem* (Chap. 10), there are large differences in the reduction of sites of the electron-transport-chain between NAD and coenzyme Q, between cytochrome *b* and cytochrome *c*, and between cytochrome *c* and cytochrome *a*. Therefore, the absence of ADP must be inhibiting electron transport at these points; in fact, these are the sites of proton extrusion leading to ATP synthesis during electron transport.

Supplementary Problems

14.11. How many molecules of ATP are produced per molecule of (*a*) pyruvate, (*b*) NADH, (*c*) glucose, and (*d*) phosphoenolpyruvate, by a cell homogenate in which glycolysis, the citric acid cycle, and oxidative phosphorylation are all completely active?

14.12. What effect would 2,4-dinitrophenol have on the P/O ratio for ATP synthesis using NADH as an electron donor?

14.13. Given that the internal pH of a mitochondrion is 7.8, and assuming that a mitochondrion has a volume the same as a sphere of diameter 1.4×10^{-6} m, calculate the number of protons inside the mitochondrion.

14.14. Why should the ingestion of an uncoupler, such as 2,4-dinitrophenol, lead to sweating, an increase in body temperature, and, in the long term (weeks), weight loss? (Note: this is an extremely dangerous method of weight control.)

14.15. How is it possible to distinguish between the different cytochromes in the electron-transport chain?

14.16. Fe^{3+} in iron-sulfur proteins has an electron spin resonance (esr) signal, while Fe^{2+} does not. Assume that you have a preparation of mitochondria that are able to synthesize ATP via oxidation of NADH; supplies of rotenone, antimycin A, and KCN; and access to an esr spectrometer. How could you establish which of the complexes of the electron-transport chain contain iron-sulfur proteins?

14.17. Calculate $\Delta G^{\circ\prime}$ for electron flow from reduced cytochrome *c* to oxygen.

14.18. Nitrite is an effective antidote to cyanide poisoning. Why? (*Hint*: Nitrite is an oxidizing agent that is able to convert Fe^{2+} to Fe^{3+} in heme groups.)

14.19. Addition of oxygen to cells metabolizing glucose under anaerobic conditions leads to (*a*) a decrease in the rate of glucose consumption and (*b*) cessation of lactate accumulation. The latter phenomenon is known as the *Pasteur effect*. Explain why these changes occur in glucose and lactate metabolism.

Chapter 15

Nitrogen Metabolism

15.1 SYNTHESIS AND DIETARY SOURCES OF AMINO ACIDS

Animals are dependent for growth on a source of *fixed* (i.e., reduced) nitrogen from other animals or plants; plants in turn are dependent on bacteria for fixing nitrogen. Humans need fixed nitrogen, which must come from the diet (normally as protein), particularly for protein and nucleic acid synthesis but also for synthesizing many specialized metabolites such as porphyrins and phospholipids.

The *amount* of protein (or fixed nitrogen) we ingest determines our state of *nitrogen balance*. Humans, like other animals, will excrete nitrogenous compounds even when they are fed on a protein-free diet, because not all nitrogenous compounds can be recycled. They are then in *negative nitrogen balance*. The amount of protein an adult needs in order to stay in nitrogen balance is not easy to define, because not all amino acids found in proteins, particularly plant proteins, are equally important for animal metabolism.

EXAMPLE 15.1

Why are plant proteins not as useful as animal protein for dietary purposes?

Cereal proteins are only about 70 percent efficient for dietary replacement purposes. The reason is that cereal proteins are deficient in lysine, an *essential amino acid* for humans (see "Amino Acid Synthesis" in this section). Thus a diet based on one source of protein (e.g., corn) can lead to malnutrition. A partial solution to the problem has been the breeding of high-lysine corn. Other plant proteins, particularly those from pod seeds (e.g., peas and beans) are deficient in the sulfur-containing amino acids. A successful vegetarian diet will therefore be balanced in cereals and pod seeds.

The accepted amount of protein required to maintain nitrogen balance is 28 g per day for a 70-kg person, i.e., about 3.8 g of nitrogen. This is estimated by measuring the N excretion over 6 to 7 days on a protein-free diet. If the protein source is from cereal, then the daily intake would have to be increased to about 40 g per day for a 70-kg person. The difference is due to the variable amounts of essential amino acids found in proteins. The amount required by growing children is larger; the accepted figure is about 0.6 g per kilogram per day.

Question: Is too much protein in the diet harmful?

Eskimos, who have a high protein intake, do have a shorter life span than Europeans, but there is no clear correlation of this with dietary protein. They also have a high intake of saturated fat.

Nitrogen Fixation

The fixation of nitrogen is the most fundamental biochemical process after photosynthesis. It is the process whereby *atmospheric nitrogen* is reduced to *ammonia*. Nitrogen fixation can be carried out by blue-green algae, some yeasts, and especially bacteria. The reduction of nitrogen

$$N_2 + 3H_2 \rightarrow 2NH_3 \qquad \Delta G^{\circ\prime} = -33.5 \text{ kJ mol}^{-1}$$

is an *exergonic* reaction. Due to the chemical unreactivity of N_2, this process is accomplished industrially by using efficient catalysts, high temperatures (600°C) and pressures (1,000 atm). The biological process occurs at 1 atm and ~25°C. In bacterial systems, the reaction is catalyzed by the enzyme *nitrogenase*.

EXAMPLE 15.2

What is the source of hydrogen and how is the chemical unreactivity of N_2 overcome in biological systems?

Protons are used as a source of hydrogen and the hydrolysis of ATP is presumed to overcome the chemical unreactivity of N_2 by assisting in the formation of thermodynamically unfavored intermediates. Eight electrons are required for the reduction of nitrogen, which is always accompanied by the evolution of H_2.

$$N_2 + 8e^- + 8H^+ \xrightarrow{16(ATP + H_2O)\ \ 16(ADP + P_i)} 2NH_3 + H_2$$

The electrons can be supplied by several donors, including NADH, flavoproteins, and NADPH.

Ultimately all higher organisms are dependent on bacterially produced ammonia for their nitrogen metabolism.

Question: How do higher organisms obtain ammonia?

Many plants, particularly legumes (peas and beans), have a symbiotic relationship with nitrogen-fixing bacteria, which live in special nodules on the roots. There are ~13,000 species of leguminous plants, all of which have symbiotic bacteria of the genus *Rhizobium*. Some insects (termites and cockroaches) also have symbiotic nitrogen-fixing bacteria in their intestines.

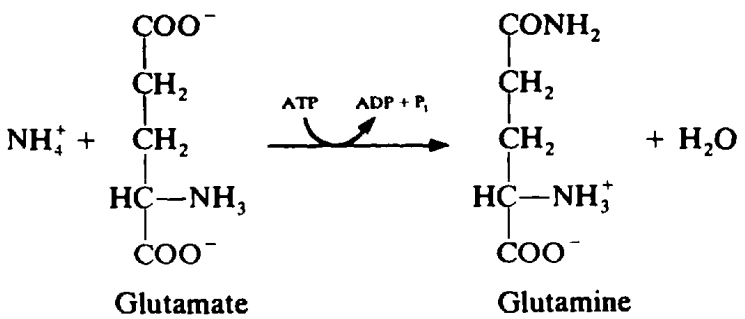
Assimilation of Ammonia

Ammonia is normally condensed with 2-oxoglutarate and thus converted to glutamate via the enzyme *glutamate dehydrogenase*; this enzyme is of highest activity in the liver and kidney. Glutamate is produced from 2-oxoglutarate and ammonia as follows:

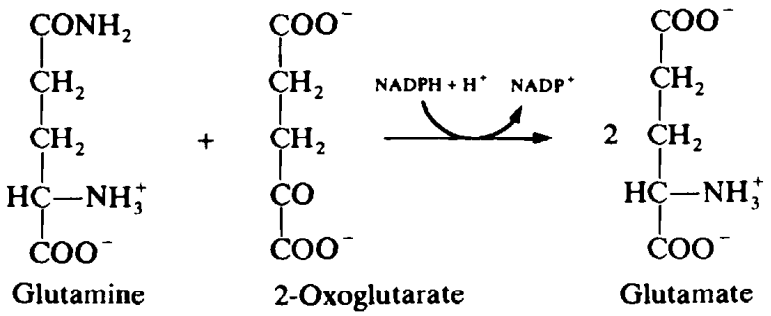
$$NH_4^+ + \text{2-Oxoglutarate}^{2-} + NADPH + H^+ \rightleftharpoons \text{Glutamate}^- + NADP^+ + H_2O$$

Glutamate dehydrogenase can also use NAD^+ for the degradation of glutamate. This reaction is freely reversible: the direction of *net flux* is determined solely by the relative concentrations of the reactants. Thus, this reaction has two equally important functions: the assimilation of ammonia or its removal from metabolites.

Glutamate is also produced in some bacteria via reactions catalyzed by the enzymes *glutamine synthetase* and *glutamate synthetase* acting together. Glutamine synthetase, as the name implies, catalyzes the synthesis of glutamine, in virtually all organisms. In humans, it is particularly active in the liver; glutamine is transported from the liver to other tissues via the blood.



Glutamate synthetase, which is not present in humans but is found in bacteria, catalyzes the formation of glutamate:



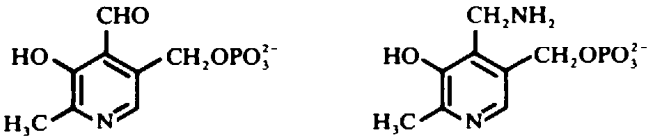
This coupled enzyme system is used by the blue-green algae and by *Rhizobia*.

The amide group of glutamine provides the ammonia for the synthesis of many N-containing compounds, e.g., purines and pyrimidines (Sec. 15.6).

Glutamate provides the amino group for the synthesis of many other amino acids through *transamination* reactions in all cells. These amino acids are then used for protein synthesis and other aspects of nitrogen metabolism. The majority of animals are dependent on plant or animal proteins for fixed nitrogen, for their nitrogen metabolism.

Transamination

Transamination, the process whereby ammonia is reversibly transferred between amino acids and 2-oxoacids, is catalyzed by *aminotransferases*, which bind *pyridoxal phosphate* as a prosthetic group. Pyridoxal phosphate and *pyridoxamine phosphate* are the coenzyme forms of vitamin B_6 (Fig. 15-1).

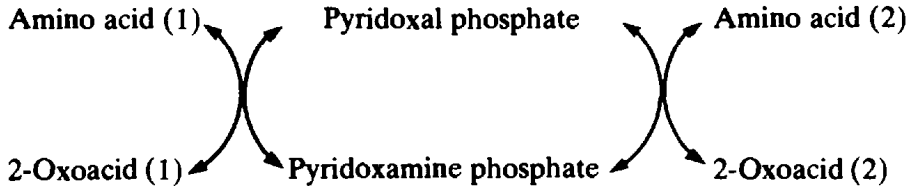


(*a*) Pyridoxal phosphate (*b*) Pyridoxamine phosphate

Fig. 15-1 Structures of pyridoxal phosphate and pyridoxamine phosphate.

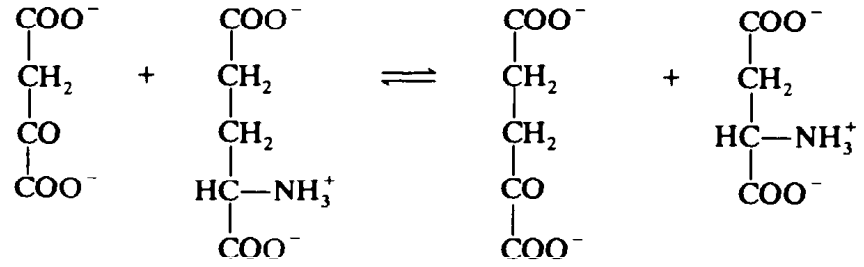
Question: What is the role of the coenzyme in transamination?

The aldehyde group of pyridoxal phosphate accepts the amino group from an amino acid by formation of a Schiff base (Chap. 1). In this process the amino acid is converted to a 2-oxoacid, and pyridoxal phosphate is converted to pyridoxamine phosphate. The amino group on pyridoxamine phosphate can now be transferred to another 2-oxoacid, converting it to an amino acid. In this second reaction, the pyridoxamine phosphate is converted back to pyridoxal phosphate.



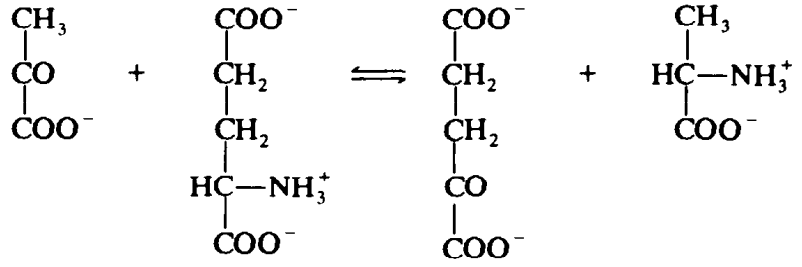
The overall reaction is: Amino acid (1) + 2-Oxoacid (2) $\rightleftharpoons$ 2-Oxoacid (1) + Amino acid (2)

2-Oxoglutarate is the normal acceptor of the amino group. In the aminotransferase reaction, 2-oxoglutarate is transaminated to give glutamate. There are at least 13 different *aminotransferases*, but their specificities are not all known. The most important are (*a*) *aspartate aminotransferase*, which catalyzes the following reversible reaction:



Oxaloacetate Glutamate 2-Oxoglutarate Aspartate

and (*b*) *alanine aminotransferase*, which catalyzes the following reaction:



Pyruvate Glutamate 2-Oxoglutarate Alanine

EXAMPLE 15.3

Both aspartate aminotransferase and alanine aminotransferase are released into the blood after damage to tissues or after cell death. Consequently, they are used as diagnostic tools when heart or liver damage has occurred, such as after a heart attack or in hepatitis, respectively. Other enzymes are also released into the blood at such times. For example, damage to heart muscle is further characterized by the presence of isoenzymes of *creatine kinase* or lactate dehydrogenase in the plasma.

The metabolic significance of the diversity of aminotransferases is not totally understood.

The highest concentration of the aminotransferases is in the cytoplasm, but they are also located in mitochondria, where *glutamate dehydrogenase* is located exclusively. The aminotransferases and glutamate dehydrogenase catalyze central reactions in amino acid metabolism. The major aminotransferases and glutamate dehydrogenase are present in all tissues in relatively high concentrations compared with other enzymes, such as those involved in glycolysis. The reversible nature of both reactions allows a rapid exchange of amino groups and formation of 2-oxoacids, as shown in Fig. 15-2.

Question: How does the scheme in Fig. 15-2 operate in starvation?

In extreme starvation, both protein and carbohydrate are in short supply. The *net* effect is that endogenous proteins (from the muscles) are hydrolyzed, releasing amino acids for protein synthesis and for oxidation to yield energy. The 2-oxoacids produced by the aminotransferases either enter gluconeogenesis (Chap. 11) or are metabolized to CO_2 and H_2O (Chap. 12). Glutamate

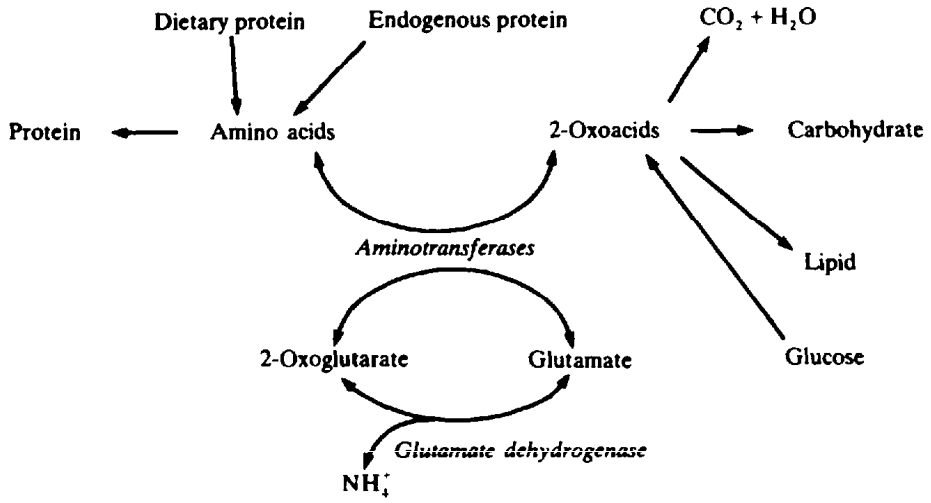


Fig. 15-2 Central role of the *aminotransferases* and *glutamate dehydrogenase* in nitrogen metabolism.

dehydrogenase, again, catalyzes the formation of ammonia from amino groups as the amino acids are broken down.

Amino Acid Synthesis

RNA base triplets (codons) exist for the 20 amino acids used in protein synthesis (Chap. 17). The ability of an organism to live and grow is dependent on protein synthesis and hence on a supply of all 20 amino acids. Higher plants are able to synthesize all 20, but many microorganisms and higher animals make considerably fewer. Humans make 10 of the 20 amino acids; the remainder must be supplied in the diet, usually in the form of plant or animal protein. The amino acids that humans cannot synthesize de novo but are essential for life are termed the *essential* amino acids. Those that we can synthesize are called the *nonessential* amino acids. The essential and nonessential amino acids are listed in Table 15.1.

Table 15.1. Nonessential and Essential Amino Acids for Humans

Nonessential	Essential
Glutamate	Isoleucine
Glutamine	Leucine
Proline	Lysine
Aspartate	Methionine
Asparagine	Phenylalanine
Alanine	Threonine
Glycine	Tryptophan
Serine	Valine
Tyrosine	Arginine*
Cysteine	Histidine

*Essential only in infants and children.

Question: What are the nonessential amino acids synthesized from?

Their syntheses depend on the availability of the appropriate *carbon skeletons* and a source of *ammonia*. Glucose is ultimately the source of the carbon skeletons for most of the nonessential amino acids. Two essential amino acids, *phenylalanine* and *methionine*, are used to form the nonessential amino acids *tyrosine* and *cysteine*, respectively. Since ammonia is available in the fed state, amino acids become essential to our diet when we are not able to synthesize their carbon skeletons.

Synthesis of 2-Oxoacids for Amino Acids

Certain 2-oxoacids are necessary for the synthesis of the nonessential amino acids; they are listed in Table 15.2.

Four of the amino acids, alanine, aspartate, glutamate, and serine, are formed by the transamination of their corresponding oxoacids. The other nonessential amino acids are then derived from these four amino acids. The syntheses of serine and tyrosine are described below because of either their importance in aspects of metabolism or their clinical significance; the synthesis of serine is essential for folic acid metabolism, while deficiencies in the enzymes synthesizing tyrosine can lead to phenylketonuria.

Table 15.2. 2-Oxoacids Required for Synthesis of the Nonessential Amino Acids

2-Oxoacid	Amino Acids
Pyruvate	Alanine
Oxaloacetate	Aspartate, asparagine
2-Oxoglutarate	Glutamate, glutamine, proline, arginine*
Pyruvate, 3-hydroxypyruvate	Serine

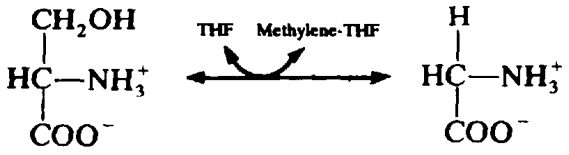
*Essential only in infants and children.

Arginine

Arginine is synthesized from aspartate and ornithine during urea formation. *Argininosuccinate synthetase* and *-lyase* catalyze the condensation and cleavage reactions, respectively, that result in the formation of arginine (Sec. 15.5).

Serine

Serine is formed from 3-phosphoglycerate (Fig. 15-3). Serine is also synthesized from glycine in a reaction catalyzed by *serine hydroxymethyltransferase*:



Serine Glycine

N^5,N^{10}-Methylenetetrahydrofolate (methylene-THF) is one of the *folic acid* coenzymes (Sec. 15.7). Note that this reaction is readily reversible, and in fact, net flux is usually in the direction of *glycine* synthesis. Thus this amino acid can arise from glucose, and it does so via serine.

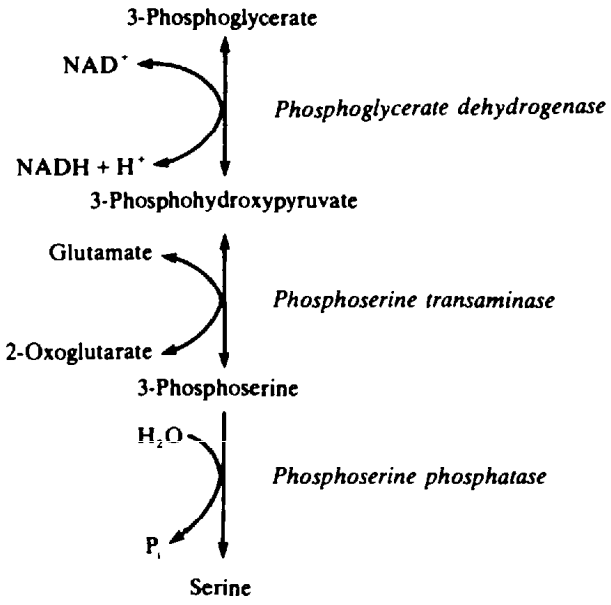


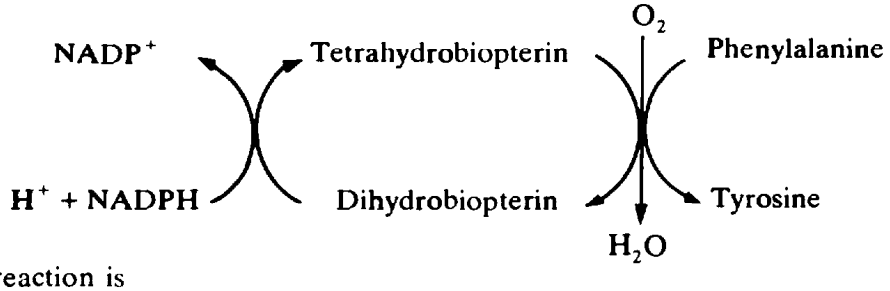
Fig. 15-3 Major pathway for serine synthesis.

Synthesis of Tyrosine and Cysteine

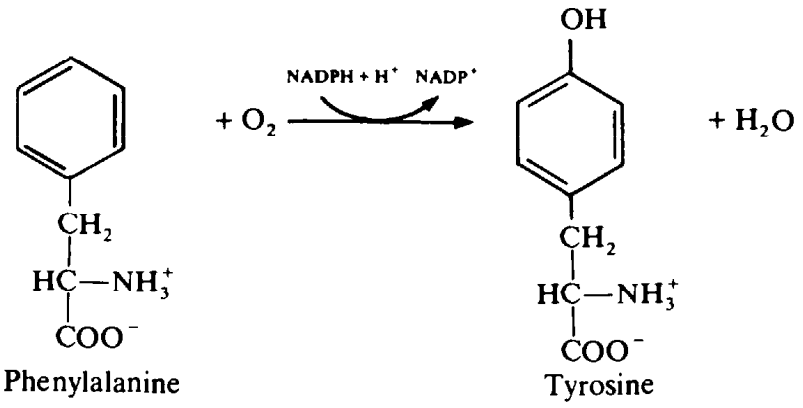
Two of the nonessential amino acids, tyrosine and cysteine, are derived from essential amino acids and may be considered to be breakdown products, as they are intermediates in the normal degradation of these amino acids. Provided sufficient of the two essential amino acids phenylalanine and methionine are available through the diet, *net synthesis* of tyrosine and cysteine can occur.

Tyrosine

Tyrosine is synthesized from phenylalanine in a reaction catalyzed by *phenylalanine hydroxylase*, which catalyzes two reactions. The reducing power in the reaction comes from NADPH, and the oxygen from molecular oxygen.



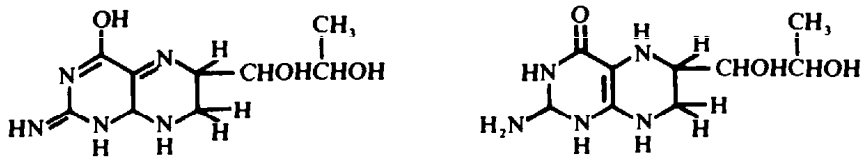
The overall reaction is



EXAMPLE 15.4

The first enzyme activity (*dihydrobiopterin reductase*) catalyzes the transfer of hydrogen to *dihydrobiopterin*, which is thus reduced to *tetrahydrobiopterin*. The second enzyme activity is a *hydroxylase* containing two Fe^{3+} atoms, and this catalyzes the reduction of O_2 such that one oxygen atom is incorporated into phenylalanine to form tyrosine and the second into water. At the same time tetrahydrobiopterin is oxidized to dihydrobiopterin. Phenylalanine hydroxylase is an example of a *mixed-function oxidase*. An inherited deficiency of phenylalanine hydroxylase results in the accumulation of phenylalanine that is not converted to tyrosine but is excreted as *phenylpyruvate*. This condition, which affects young infants, is known as *phenylketonuria* and is associated with severe mental retardation.

Biopterin, like folic acid (Sec. 15.7, Fig. 15-19), contains the pterin ring (Fig. 15-4).



(*a*) Dihydrobiopterin (*b*) Tetrahydrobiopterin

Fig. 15-4 Structures of biopterin derivatives.

15.2 DIGESTION OF PROTEINS

Dietary protein is the principal source of fixed nitrogen in higher animals. In digestion, proteins are hydrolyzed by a series of hydrolytic enzymes in the stomach and the small intestine to peptides and amino acids, which are absorbed from the lumen of the gastrointestinal tract. These enzymes are known collectively as *proteolytic enzymes*, or *proteases*, and belong to the class of enzymes called *hydrolases* (Chap. 8).

Zymogens

The proteolytic enzymes are secreted in the gastric juice or by the pancreas as inactive precursors called *zymogens*. In the case of *trypsin*, the zymogen, *trypsinogen*, is synthesized on the endoplasmic reticulum of pancreatic cells and is secreted from *zymogen granules* into a duct that leads to the duodenum. The granules are produced in the Golgi apparatus and consist of trypsinogen molecules surrounded by a lipid-protein membrane. The pancreatic cells also produce a *trypsin inhibitor* that ensures that they are not autodigested.

EXAMPLE 15.5

In the disease *pancreatitis*, that occasionally follows a bout of mumps, the proteolytic enzymes secreted by the pancreas are prematurely activated and digest the cells of the pancreas.

The entry of protein into the stomach stimulates the release of a hormone, *gastrin*, which then causes the release of hydrochloric acid from the *parietal cells*, and *pepsinogen* from the *chief cells* (Fig. 15-5). Pepsinogen is another zymogen (they all start with *pro-* or end in *-ogen*) that is converted in the gastric juice to the active enzyme *pepsin*.

Question: What is the function of hydrochloric acid in digestion?

The hydrochloric acid lowers the pH of the stomach contents to ~pH2, which kills most microorganisms and denatures proteins, making their peptide bonds more accessible to enzymatic hydrolysis.

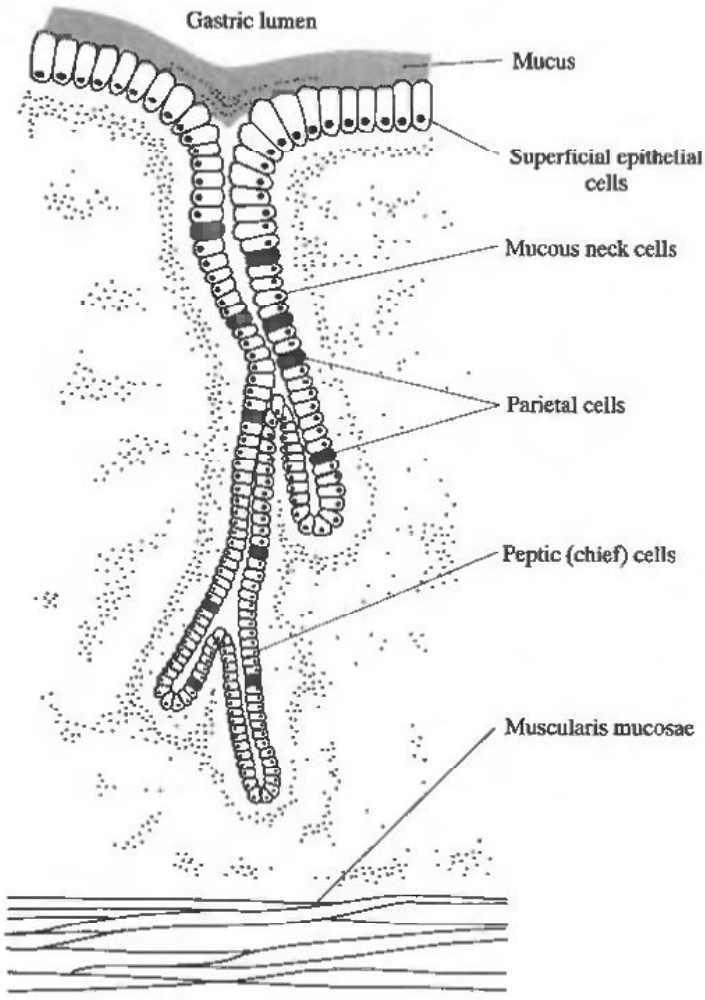


Fig. 15-5 Schematic portrayal of a gastric gland.

When the stomach contents pass into the small intestine, the low pH causes the release of the hormone *secretin* from cells of the small intestine. Secretin causes the release of bicarbonate from the pancreas which neutralizes the hydrochloric acid and allows the hydrolytic enzymes *trypsin*, *chymotrypsin*, *elastase* and *carboxypeptidase* to function optimally at pH 7–8.

EXAMPLE 15.6

There are a variety of *peptide hormones* acting in the gut: the *gastrins* stimulate gastric acid secretion; *secretin* and *somatostatin* inhibit the production of gastrins. *Cholecystokinin* and *somatostatin* can inhibit gastric acid secretion directly, and the former one causes the gall-bladder to contract and thus force bile into the duodenum.

In the duodenum, the pancreatic zymogens, *trypsinogen*, *chymotrypsinogen*, *proelastase* and *procarboxypeptidase* are converted into active enzymes by enteropeptidase and trypsin, as shown in Fig. 15-6. The activation of all the zymogens involves cleavage of peptide bonds and removal of peptides, enabling conformational changes and formation of a functional active site.

Question: What is the actual sequence of chemical events involved in zymogen activation?

Trypsinogen, chymotrypsinogen, proelastase, and procarboxypeptidase are all synthesized as single polypeptide chains with an M_r around 25,000–30,000. The initial step in the activation is the

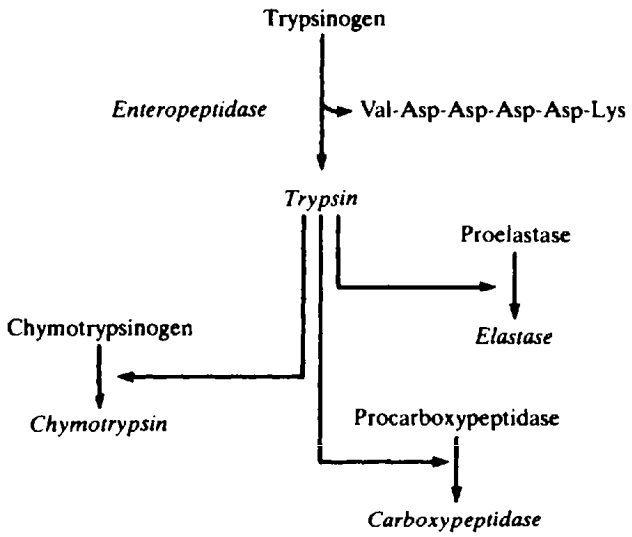


Fig. 15-6 Activation of pancreatic zymogens.

hydrolysis of a hexapeptide from the N terminus of trypsinogen. This hydrolysis produces trypsin and is catalyzed by *enteropeptidase*, a large glycoprotein enzyme situated on the membranes of brush border cells of the small intestine.

Question: How are zymogens other than trypsinogen activated?

The activation has been studied in detail; that of chymotrypsinogen is shown in Fig. 15-7.

Chymotrypsinogen, a single polypeptide chain of 245 amino acid residues, is converted to α-chymotrypsin, which has three polypeptide chains linked by two of the five disulfide bonds present in the *primary* structure of chymotrypsinogen. π- and δ-chymotrypsin also have proteolytic activity. In contrast, the conversion of *procarboxypeptidase* to *carboxypeptidase* involves the hydrolytic removal of a single amino acid.

Specificity of Proteases

Theoretically, there are 20×20 possible different combinations of amino acid residues adjacent to one another in a polypeptide. If each possible combination were needed a specific protease, then 400 different proteolytic enzymes would be required. However, the proteolytic enzymes have broad specificities, largely confined to groups of amino acids with similar side-chain characteristics (e.g., basic or nonpolar), and therefore only a few different types of enzymes are encountered.

Question: What determines the substrate specificities of the proteolytic enzymes?

All the proteolytic enzymes catalyze the hydrolysis of peptide bonds:

$$R{-}CO{-}NH{-}R' + H_2O \rightleftharpoons R{-}COO^- + NH_3^+{-}R'$$

The specificities are determined by the side chains of the amino acids on either side of the peptide bond that is hydrolyzed in the polypeptide chain. For the endopeptidases, it is the side chain of the amino acid contributing the carbonyl group of the peptide bond that determines whether the substrate will bind. Thus, chymotrypsin hydrolyzes peptide bonds where the carbonyl group is from one of the *aromatic* amino acids, namely, phenylalanine, tyrosine, or tryptophan. The specificities are listed in Table 15.3.

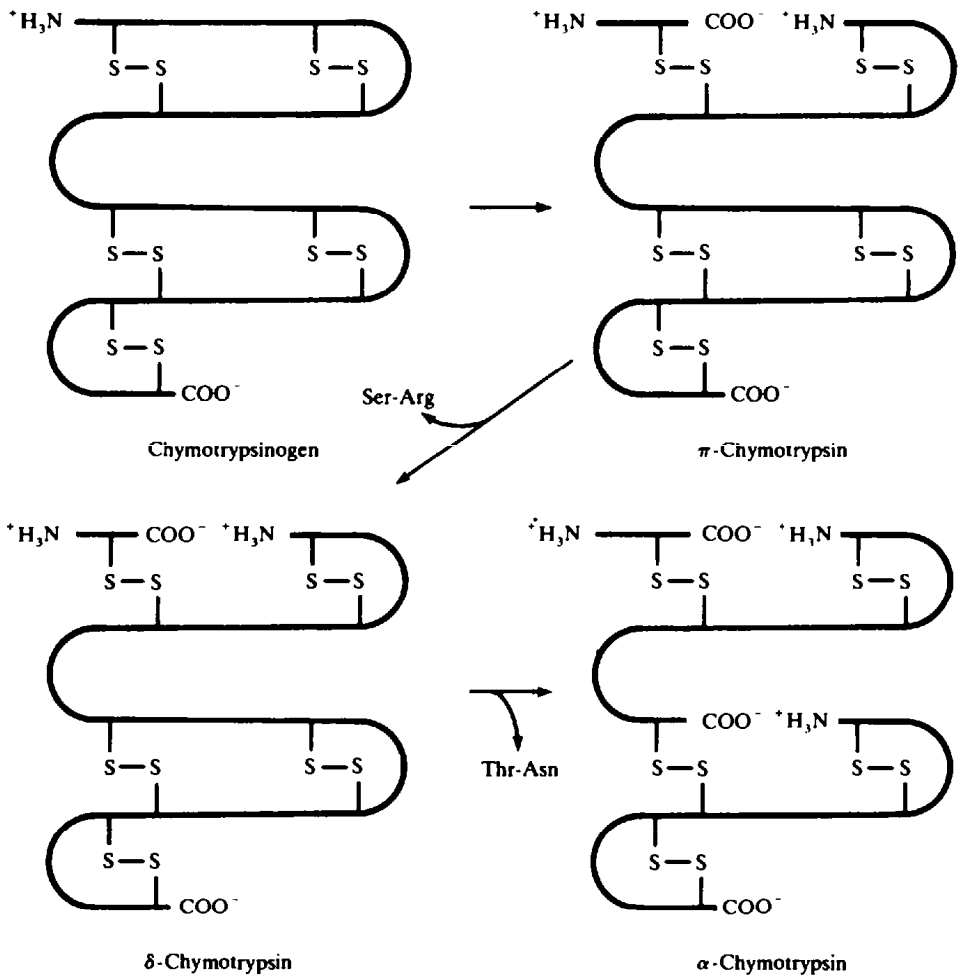


Fig. 15-7 Activation of chymotrypsinogen.

Question: Are there any similarities in the mechanisms of catalysis of the pancreatic proteases?

Three of the four pancreatic proteases (trypsin, chymotrypsin, and elastase) are called *serine proteases* because they are all dependent for activity on the side chain of a serine residue in the active site. This serine residue attacks the carbonyl group of the peptide bond to cleave the peptide, giving an acyl-enzyme intermediate (Chap. 8). This ester bond is then hydrolyzed in a second step:

$$R{-}CO{-}NH{-}R' + Enz{-}CH_2OH \rightleftharpoons R{-}CO{-}OCH_2{-}Enz + R'NH_2$$

$$R{-}CO{-}OCH_2{-}Enz + H_2O \rightarrow R{-}COOH + Enz{-}CH_2OH$$

Table 15.3. Specificities of Proteolytic Enzymes

Enzyme	Specificity
Pepsin	Phe, Tyr, Trp; also, Leu, Glu, Gln
Trypsin	Lys, Arg
Chymotrypsin	Phe, Tyr, Trp
Carboxypeptidase A	A bulky, nonpolar carboxy-terminal residue
Elastase	Ala, Gly, Ser
Aminopeptidase	Any amino-terminal residue

EXAMPLE 15.7

The different specificities of the proteolytic enzymes are due to *specificity pockets* at the binding site (Fig. 15-8). These pockets on the surface of the enzyme accommodate the side-chain of the amino acid residue located on the carbonyl side of the scissile bond of the substrate. In trypsin, a serine residue present in chymotrypsin is replaced by an aspartate residue. This allows the binding of cationic arginine and lysine residues instead of bulky aromatic side chains. In elastase, two glycine residues of chymotrypsin are replaced by valine and threonine. Their bulky side chains block the specificity pocket so that elastase hydrolyzes peptide bonds adjacent to smaller, uncharged side chains.

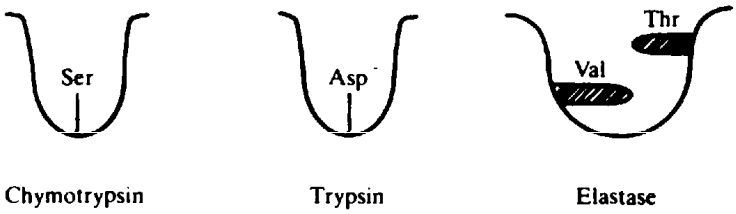


Fig. 15-8 Substrate specificity pockets.

Pepsin and the pancreatic proteases catalyze the conversion of dietary protein to peptides and amino acids. The *aminopeptidases* and the *dipeptidases* in the intestinal mucosa almost complete the hydrolysis of the peptides to amino acids, but some peptides, especially those containing glutamate, pass into the gut mucosal cells with the free amino acids. The aminopeptidases remove amino acids from the N-terminus of a peptide.

The hydrolysis of proteins in digestion is summarized in Fig. 15-9.

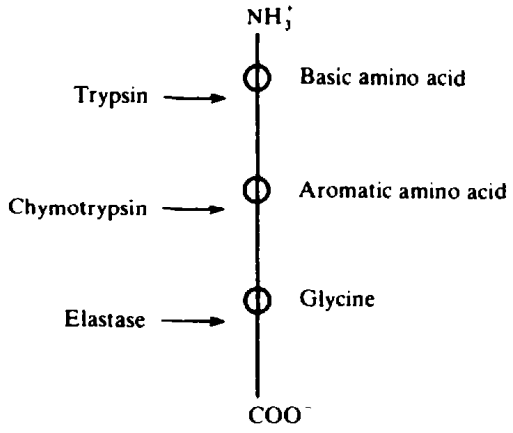


Fig. 15-9 Breakdown of proteins to amino acids in digestion.

Amino Acid Transport

Amino acids, dipeptides, and some tripeptides are transported from the lumen of the intestine through the membrane of the brush border of the mucosal cells and into the cytoplasm, where the peptides are hydrolyzed to amino acids. Transport of peptides and amino acids is *active* and analogous to glucose transport; i.e., they are transported, together with Na^+, across the gut-cell membrane by specific transport proteins called *Na^+ symports*. Between the gut lumen and the cytoplasm of the cell there is a concentration gradient of Na^+ that is maintained by Na^+/K^+ ATPase at the base of the cell adjacent to the blood capillaries; this Na^+/K ATPase pumps Na^+ from the cell into the blood.

Thus, the Na^+ concentration is lower in the cell than in the gut lumen, ensuring that the amino acids and peptides will be transported across the membrane as Na^+ moves in. There are at least seven different transport proteins involved in the uptake of the amino acids.

EXAMPLE 15.8

How do seven transport proteins operate on 20 different amino acids?

The transporters have overlapping specificities. Thus, there is one transporter (called *system L*) for leucine and neutral amino acids with branched or aromatic side chains, another for basic amino acids (the *Ly system*), and a low-activity carrier (the *dicarboxylate system*) for dicarboxylic amino acids.

Some of the amino acids undergo facilitated diffusion through selective transport proteins into the bloodstream, from which they are taken up by the liver and other organs. Others, particularly glutamate, glutamine, aspartate, and asparagine, are metabolized by the gut cells for energy.

15.3 DYNAMICS OF AMINO ACID METABOLISM

In addition to being synthesized or produced by the hydrolysis of dietary protein, amino acids can come from hydrolysis of tissue proteins, e.g., intestinal mucosa or, during starvation, muscle. Amino acids are used in protein synthesis (Chap. 17); they also enter gluconeogenesis and lipogenesis; are degraded to provide energy; and are used for synthesizing compounds such as purines, pyrimidines, porphyrins, epinephrine and creatine.

This metabolic activity is achieved by a turnover of amino acids and proteins that is as rapid as that of lipids and carbohydrates. In an adult human male, 400 g of body proteins is turned over each day. Of this, 50 g is used to replace digestive enzymes (Sec. 15.2), and 6 g to replace hemoglobin (Sec. 15.8). The concentration of free amino acids in plasma is small (total ~3.2 mmol L^{-1}, of which 25 percent is glutamine), but the turnover of 400 g per day of protein is equivalent to the uptake, and release back into the plasma, of 4.6 moles of α-amino-N, so that the average lifetime of an amino

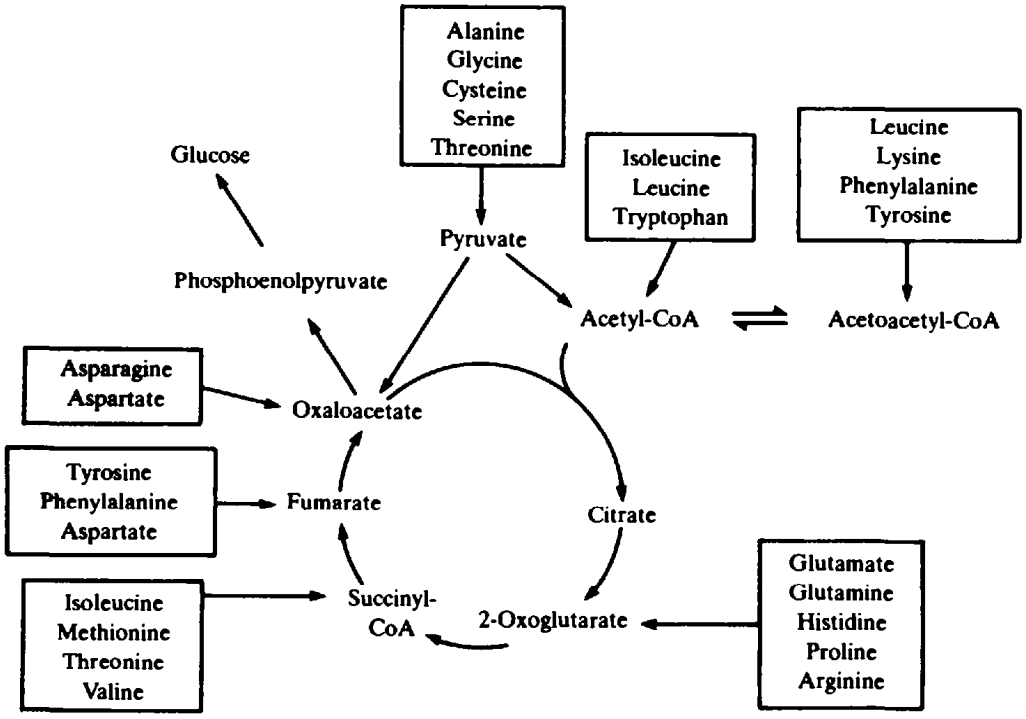


Fig. 15-10 Fates of the carbon skeletons of amino acids.

acid in the plasma is about 5 min. Plasma amino acids are turned over with the same kind of rapidity as plasma glucose or free fatty acids. Like that of plasma glucose, the plasma amino acid concentration is remarkably constant, but it is not understood how this is maintained.

15.4 AMINO ACID CATABOLISM

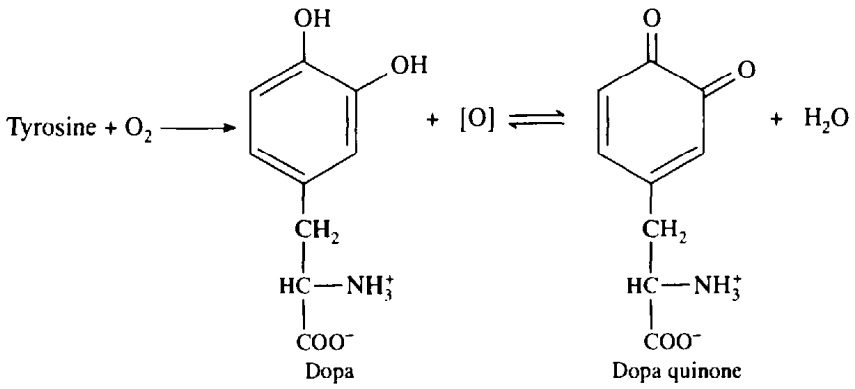
The catabolism of the amino acids is complex; there are too many differences between the amino acids for any useful generalizations to be made.

The carbon skeletons of the amino acids, with the exception of leucine, can be used for gluconeogenesis. The fates of the carbon atoms are summarized in Fig. 15-10.

All those amino acids which can eventually give rise to pyruvate, which in turn can be used for gluconeogenesis, are called *glucogenic amino acids*. The one amino acid, leucine, that does not give any intermediates of gluconeogenesis (i.e., it is only ketogenic) is degraded to acetoacetate and acetyl-CoA. Several amino acids, phenylalanine, tyrosine, tryptophan, and isoleucine, are both glucogenic and ketogenic. Thus the *majority* of the amino acids are at best glucogenic. The pathways involved in the catabolism of the individual amino acids range from one-step reactions, such as with aspartate, glutamate, and alanine, which use the appropriate aminotransferases, to multistep pathways, as with the aromatic amino acids and lysine; e.g., tyrosine is degraded in four steps to acetoacetate and fumarate.

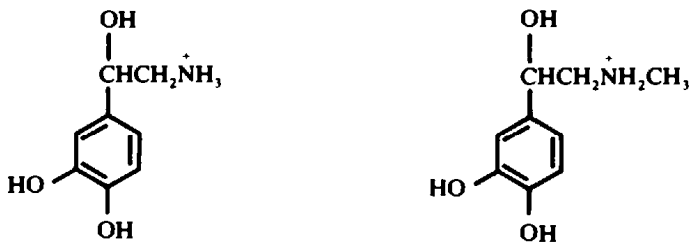
EXAMPLE 15.9

Tyrosine, itself a degradation product of phenylalanine (Sec. 15.1), is initially converted to *3,4-dihydroxyphenylalanine* (*dopa*), and the corresponding *dopa quinone*, by the copper-containing enzyme *tyrosinase*. Tyrosinase is found in *melanocytes* and is a mixed-function oxidase. It catalyzes the following reaction:



Dopa quinone is converted to *norepinephrine* and *epinephrine* (Fig. 15-11) in the adrenal medulla.

Amino acid catabolism is particularly important during starvation when the amino acids catabolized in the major tissues reflect the function of those tissues. Because of the bulk of muscle, amino acid catabolism is particularly important in this tissue which, in starvation, supplies the liver with most of its gluconeogenic precursors. Amino acids resulting from proteolysis during starvation are interconverted in the muscle so that about 60% of the amino acid mass leaving the muscle is either glutamine or alanine. The branched chain amino acids, valine, leucine and isoleucine, which are all essential amino acids, are deaminated in muscle by a specific aminotransferase and the 2-oxoacids are transported to the liver for further metabolism by branched chain 2-oxoacid dehydrogenase (BCOADH). The aminotransferase is inactive in liver which ensures that the peripheral tissues are supplied with valine, leucine and isoleucine. In times of plenty, the activity of BCOADH, which oxidatively decarboxylates all three amino acids in a reaction analogous to that



(*a*) Norepinephrine (*b*) Epinephrine

Fig. 15-11 Structures of norepinephrine and epinephrine.

catalyzed by the pyruvate dehydrogenase complex into CoA derivatives, is regulated by phosphorylation (which leads to the inactive form) in accordance with the body's demand for branched chain amino acids. This control is overriden during starvation when the body's demand for glucose, for survival, becomes paramount.

In starvation, the kidney uses glutamine, and glutamate derived from it, as a source of ammonia to buffer ketone bodies that are excreted. Some ammonia goes to the liver for urea synthesis. The

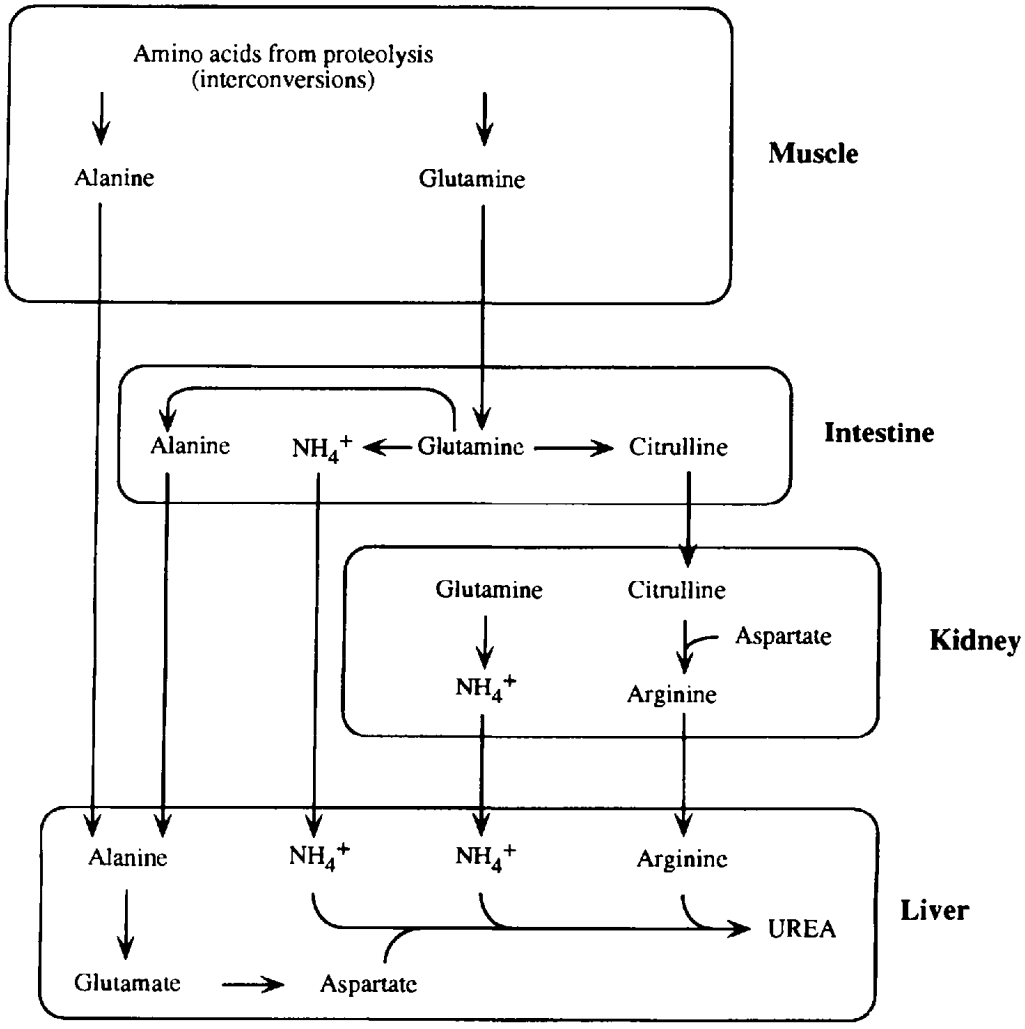


Fig. 15-12 Transfer of nitrogen components from the tissues to the liver for urea synthesis.

carbon skeleton, 2-oxoglutarate, is used for gluconeogenesis. The intestine preferentially uses glutamine, glutamate, aspartate and asparagine for metabolism in both normal times and in starvation. The intestine has a high demand for cell division due to the sloughing of intestinal cells, and glutamine is used as a source of nitrogen for purine synthesis. Some glutamine is used to make citrulline (which travels to the kidney to be converted to arginine for use later by the liver in urea synthesis) and the remainder is converted into alanine which enters the portal vein. Figure 15-12 summarizes amino acid metabolism in the tissues discussed, and the origin of nitrogen for urea synthesis in the liver.

15.5 DISPOSAL OF EXCESS NITROGEN

Question: Why is nitrogen excreted?

There are no stores of nitrogen in the body comparable in form to lipid and glycogen; i.e., any nitrogen in excess of our growth requirements is excreted. If we ingest less nitrogen than we require for normal growth and tissue repair, we utilize the nitrogen stored in our muscle proteins.

Amino acids in excess of metabolic requirements are degraded to their carbon skeletons, which, as discussed in the preceding section, enter energy metabolism or are converted to other compounds and ammonia; the ammonia is excreted as such or is converted to urea and then excreted. This is the situation with humans, but there are major species differences as to the means by which excess ammonia is removed.

In aquatic animals, ammonia diffuses out of the body through the skin, but land animals excrete excess ammonia either as urea or uric acid. Ammonia is excreted by humans on high meat diets as a strategy to conserve Na^+ and K^+. Excess PO_4^{3-} and SO_4^{2-} produced from phosphoproteins and S-containing amino acids are excreted as ammonium salts; Na^+ and K^+ are exchanged for NH_4^+ in the kidney. The excretion of urea requires a plentiful supply of water, as it is normally excreted in solution, whereas uric acid is very insoluble and is excreted as a solid by birds and reptiles. Thus, in animals in which weight, or the conservation of water, is important (e.g., birds), excess ammonia is excreted as uric acid.

Urea, NH_2CONH_2, is highly soluble (10 mol L^{-1}), nontoxic, and high in nitrogen content (47 percent). Normal human subjects excrete ~30 g per day on a western diet, but on a high-protein diet, this can increase to 100 g per day. Humans, and other primates, excrete a small amount of uric acid as an end product of purine metabolism. We excrete excess nitrogen as ammonia, urea, and uric acid. Several other nitrogen-containing metabolites, notably the bile pigments, are also excreted. These are degradation products of hemoglobin and other porphyrin-containing molecules.

Formation of Ammonia

The major enzyme involved in the formation of ammonia in the liver, brain, muscle, and kidney is *glutamate dehydrogenase*, which catalyzes the reaction in which ammonia is condensed with 2-oxoglutarate to form glutamate (Sec. 15.1). Small amounts of ammonia are produced from important amine metabolites such as epinephrine, norepinephrine, and histamine via *amine oxidase* reactions. It is also produced in the degradation of purines and pyrimidines (Sec. 15.6) and in the small intestine from the hydrolysis of glutamine. The concentration of ammonia is regulated within narrow limits; the upper limit of normal in the blood in humans is ~70 μmol L^{-1}. It is toxic to most cells at quite low concentrations; hence there are specific chemical mechanisms for its removal. The reasons for ammonia toxicity are still not understood. The activity of the urea cycle in the liver maintains the concentration of ammonia in peripheral blood at ~20 μmol L^{-1}.

In tissues, ammonia and the ammonium ion are in equilibrium:

$$NH_3 + H^+ \rightleftharpoons NH_4^+$$

At a physiological pH of 7.2, ~99 percent of the ammonia is in the ionic form. The un-ionized form diffuses across cell membranes while the ammonium ion is transported more slowly via carrier-mediated processes.

It should be noted that a large fraction of the ammonia that is converted to urea in the liver comes from metabolism in the extrahepatic tissues, though only a small fraction leaves these tissues as ammonia. The absorptive cells of the small intestine are exceptional in that they release ammonia into the portal vein; there the ammonium concentration may reach 0.26 mmol L^{-1}, accounting for 30 percent of the urea synthesized in the liver. The flow of nitrogen compounds to the liver where urea is synthesized is shown in Fig. 15-12.

Urea Synthesis

Urea is synthesized in the liver by a series of reactions known as the *urea cycle* (Fig. 15-13). One nitrogen is derived from ammonium, the second from aspartate; and the carbon is derived from CO_2. The synthesis of urea requires the formation of *carbamoyl phosphate* and the four enzymatic reactions of the urea cycle. Some of the reactions take place in the mitochondria and some in the cytoplasm. The enzymes involved in the synthesis of urea are discussed below.

Carbamoyl Phosphate Synthetase I

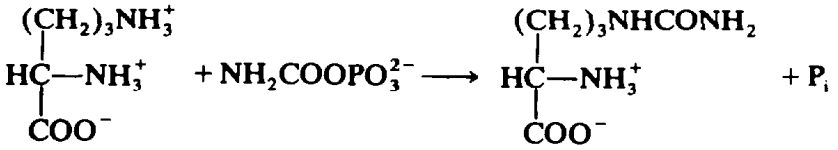
The formation of carbamoyl phosphate ($NH_2COOPO_3^{2-}$) takes place in the matrix of mitochondria:

$$NH_4^+ + HCO_3^- + 2ATP \rightarrow NH_2COOPO_3^{2-} + 2ADP + P_i$$

The ammonium can come from glutamate via glutamate dehydrogenase or in free form from the blood, and the HCO_3^- comes from respiration.

Ornithine Carbamoyltransferase

The first reaction of the urea *cycle* takes place in the matrix of the mitochondria and is catalyzed by ornithine carbamoyltransferase.



L-Ornithine Carbamoyl phosphate Citrulline

Both ornithine, which is a homologue of lysine, and citrulline are L-amino acids, but neither has a genetic codon, and both are found only as *posttranslational modifications* of arginine residues in some proteins such as keratin. Citrulline leaves the mitochondria by the same transport system that facilitates ornithine's entry from the cytoplasm.

EXAMPLE 15.10

Citrulline takes its name from the watermelon genus (*Citrullus*) in which it was first found in 1930. It was also discovered the same year as a bacterial degradation product of arginine. Krebs, who elucidated the form of the urea cycle, demonstrated that citrulline was the intermediate between ornithine and arginine. The urea cycle was the first metabolic cycle to be discovered. In Krebs' words, "it revealed a new pattern of the organization of metabolic processes."

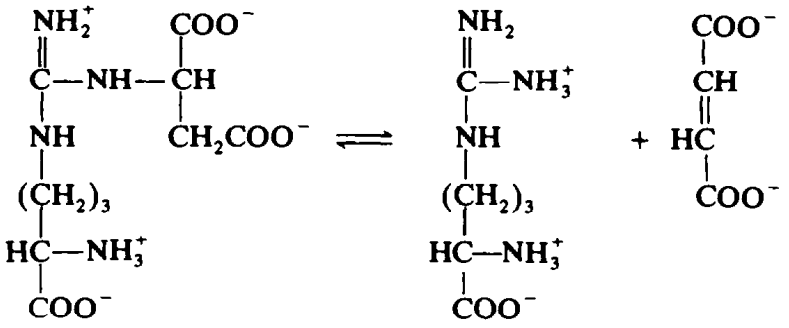
Argininosuccinate Synthetase

Argininosuccinate synthetase (the second enzyme of the urea cycle) and the remaining two enzymes of the cycle are found in the cytoplasm. Argininosuccinate synthetase catalyzes the condensation of citrulline with aspartate to form *argininosuccinate*. The reaction requires one molecule of ATP, which is hydrolyzed to AMP and PP_i. Pyrophosphate is a strong inhibitor of the reaction ($K_i = 6.2 \times 10^{-5} M$); but inhibition is normally not evident because of pyrophosphatase activity.

$$\text{Citrulline} + \text{Aspartate} + \text{ATP} \overset{Mg^{2+}}{\rightleftharpoons} \text{Argininosuccinate} + \text{AMP} + PP_i$$

Argininosuccinate Lyase

Argininosuccinate lyase (the third enzyme) reversibly catalyzes the cleavage of argininosuccinate to arginine and fumarate.



Argininosuccinate Arginine Fumarate

This reaction also supplies arginine for protein synthesis. Any arginine removed from the cycle in this way must be replaced. This is done by synthesizing ornithine from glutamate. The fate of fumarate depends on the need for gluconeogenesis. If glucose is needed, fumarate is converted via cytosolic fumarase and malate dehydrogenase to oxaloacetate and thence to phosphoenolpyruvate and glucose. If the fumarate is not needed for gluconeogenesis, it can be converted to oxaloacetate and transaminated by aspartate aminotransferase to provide aspartate for a further round of the urea cycle.

Arginase

Arginase, the last (fourth) enzyme in the urea cycle, catalyzes the hydrolytic cleavage of arginine to urea and ornithine.

$$\text{Arginine} + H_2O \rightarrow \text{Urea} + \text{Ornithine}$$

The urea passes via a transport protein into the blood and is carried to the kidneys where it enters the glomerular filtrate, from which it is excreted in the urine.

The reactions of the urea cycle are summarized in Fig. 15-13.

The overall reaction of the urea cycle is:

$$3\text{ATP} + NH_4^+ + CO_2 + 2H_2O + \text{Aspartate} \rightarrow 2\text{ADP} + 4P_i + \text{AMP} + \text{Fumarate} + \text{Urea}$$

Since the regeneration of ATP from AMP requires a molecule of ATP to convert AMP to ADP (a reaction catalyzed by the enzyme *adenylate kinase*), a total of four molecules of ATP are hydrolyzed in the synthesis of one molecule of urea.

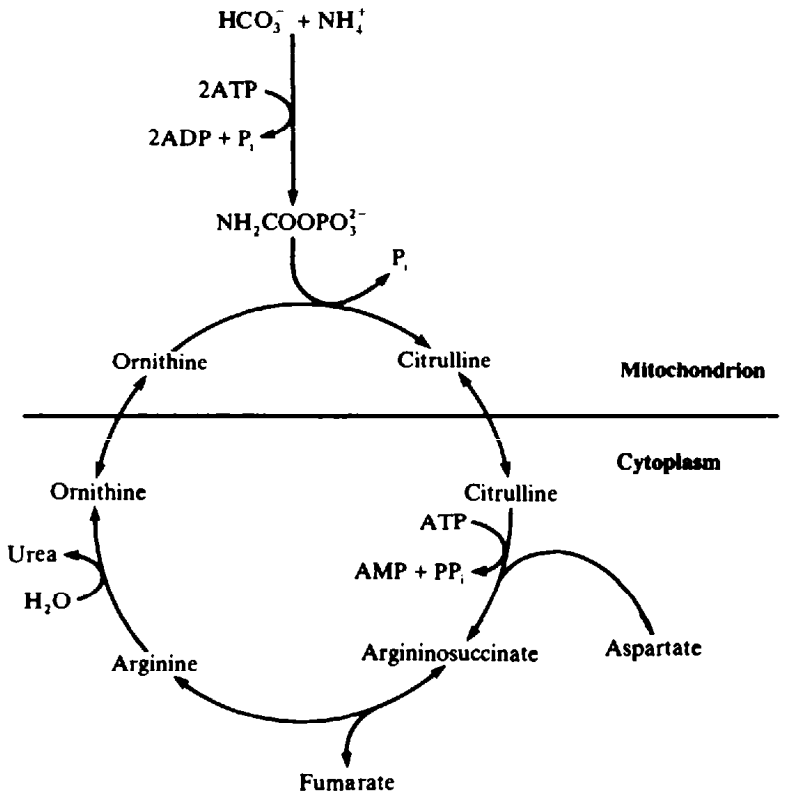


Fig. 15-13 The urea cycle.

Question: Why is the urea cycle compartmentalized?

The main reason is probably that the system evolved to keep the fumarate concentration low, because fumarate (and arginine) readily inhibits argininosuccinate lyase. Thus, this enzyme is cytoplasmic; it is not inhibited by the high concentration of fumarate from the citric acid cycle since this fumarate is in the mitochondrion.

15.6 PYRIMIDINE AND PURINE METABOLISM

The synthesis of nucleotides is important not only because of the crucial role nucleic acids play in protein synthesis and the storage of genetic information, but also because of the role nucleotides such as FAD, NAD(P)H, CoASH, cAMP and UDP-glucose play in metabolism.

Biosynthesis of Pyrimidine Nucleotides

The atoms of the pyrimidine ring are derived from *carbamoyl phosphate* and *aspartate*, as shown in Fig. 15-14. The *de novo* biosynthesis of pyrimidine nucleotides is shown in Fig. 15-15. The first completely formed pyrimidine ring is that of *dihydroorotate*. Only after oxidation to *orotate* is the ribose attached to produce orotidylate. The compound 5-phosphoribosyl 1-pyrophosphate (P-Rib-PP) provides the ribose phosphate. L-Glutamine is used as a substrate donating nitrogen atoms at reactions 1 and 9, catalyzed by *carbamoyl phosphate synthetase II* and *CTP synthetase*, respectively; a second

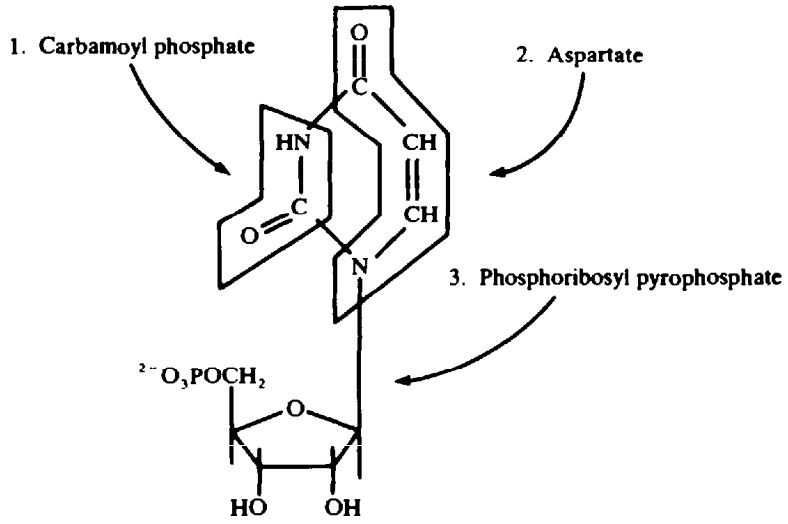


Fig. 15-14 Components in the synthesis of uridine monophosphate.

amino acid, L-aspartate, is a substrate for reaction 2, catalyzed by *aspartate transcarbamoylase*. P-Rib-PP is an activator of carbamoyl phosphate synthetase II and a substrate for reaction 5, catalyzed by *orotate phosphoribosyltransferase*.

The end-product of the pathway, UTP, is a potent inhibitor of carbamoyl phosphate synthetase II and the substrate, ATP, also activates this enzyme. The enzymatic activity of carbamoyl phosphate synthetase II is low relative to subsequent enzymes in the pathway (Fig. 15-15) and, under normal conditions, flux through the *de novo* pathway may be regulated by cellular levels of P-Rib-PP, UTP and ATP; i.e., carbamoyl phosphate synthetase II catalyzes the flux-controlling step in the pathway (see Section 10.5).

There are two multifunctional proteins in the pathway for *de novo* biosynthesis of pyrimidine nucleotides. A *trifunctional* protein, called *dihydroorotate synthetase* (or *CAD*, where the letters are the initials of the three enzymatic activities), catalyzes reactions 1, 2 and 3 of the pathway ($HCO_3^- \rightarrow$ CAP $\rightarrow$ CA-asp $\rightarrow$ DHO; Fig. 15-15). The enzymatic activities of carbamoyl phosphate synthetase, aspartate transcarbamoylase and dihydroorotase, are contained in discrete globular domains of a single polypeptide chain of 243 kDa, where they are covalently connected by segments of polypeptide chain whch are susceptible to digestion by proteases such as trypsin. A *bifunctional enzyme*, *UMP synthase*, catalyzes reactions 5 and 6 of the pyrimidine pathway (orotate $\rightarrow$ OMP $\rightarrow$ UMP; Fig. 15-15). Two enzymatic activities, those of *orotate phosphoribosyltransferase* and *OMP decarboxylase*, are contained in a single protein of 51.5 kDa which associates as a dimer.

Dihydroorotate dehydrogenase, the enzyme catalyzing the dehydrogenation of dihydroorotate to orotate (reaction 4 of the pathway; Fig. 15-15), is located on the outer side of the inner mitochondrial membrane. This enzyme has FAD as a prosthetic group and in mammals electrons are passed to ubiquinone. The *de novo* pyrimidine pathway is thus compartmentalized; dihydroorotate synthesized by trifunctional DHO synthetase in the cytosol must pass across the outer mitochondrial membrane to be oxidized to orotate, which in turn passes back to the cytosol to be a substrate for bifunctional UMP synthase. Mammalian cells contain two carbamoyl phosphate synthetases: the *glutamine-dependent enzyme* (*CPSase II*) which is part of CAD, and an *ammonia-dependent enzyme* (*CPSase I*) which is found in the mitochondrial matrix, and which is used for urea and arginine biosynthesis. Under certain conditions (e.g., hyperammonemia), carbamoyl phosphate synthesized in the matrix by CPSase I may enter pyrimidine biosynthesis in the cytosol.

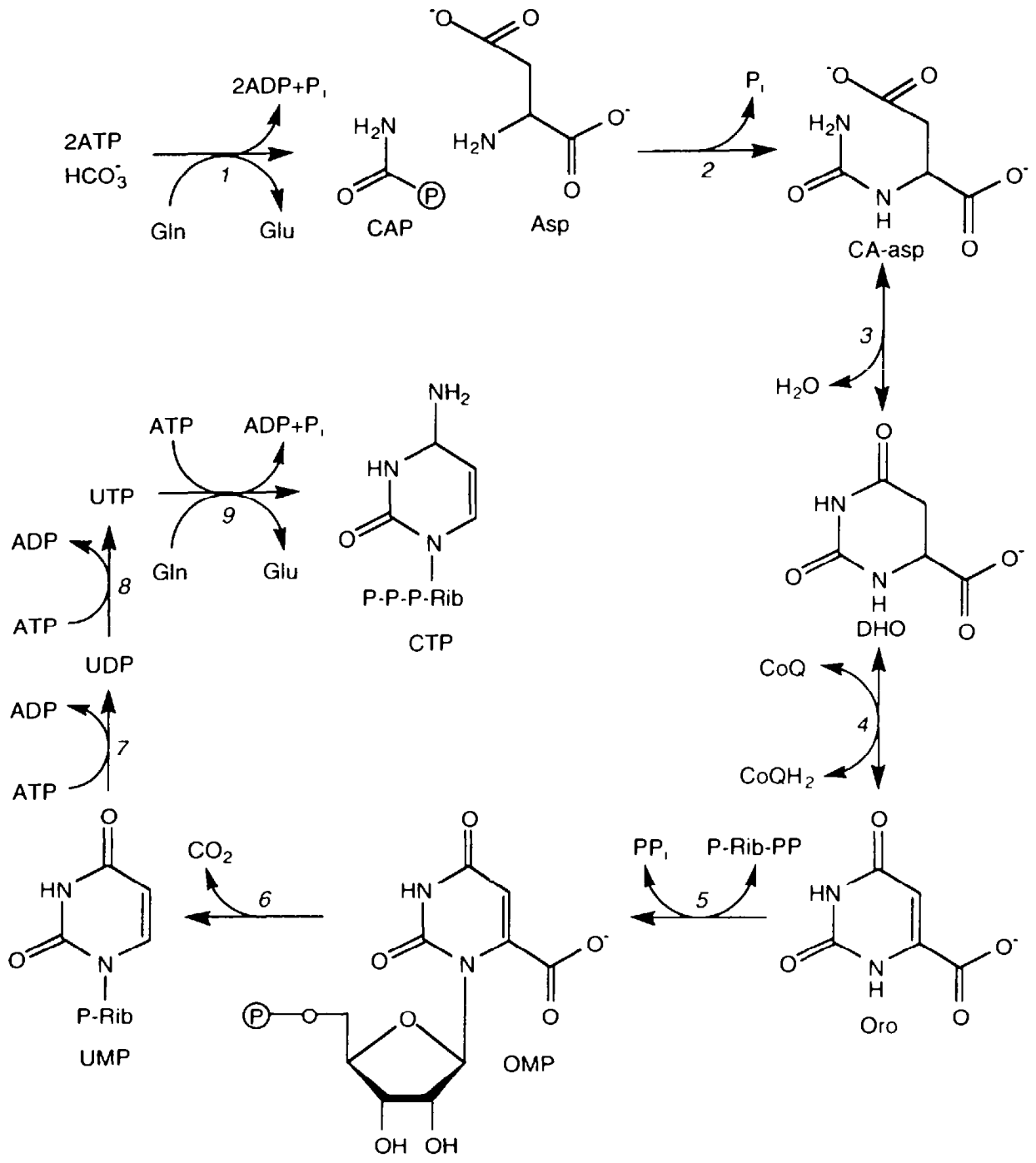


Fig. 15-15 The *de novo* pyrimidine biosynthetic pathway. CAP, carbamoyl phosphate; CA-asp, *N*-carbamoyl-L-aspartate; DHO, L-dihydroorotate; Oro, orotate; OMP, orotidine 5'-monophosphate. Enzymes: (1) carbamoyl phosphate synthetase II; (2) aspartate transcarbamoylase; (3) dihydroorotase, (4) dihydroorotate dehydrogenase; (5) orotate phosphoribosyltransferase; (6) OMP decarboxylase; (7) nucleoside monophosphate kinase; (8) nucleoside diphosphate kinase; (9) CTP synthetase.

EXAMPLE 15.11

In bacteria, such as *E. coli*, the first six reactions of the *de novo* pyrimidine pathway (Fig. 15-15) are catalyzed by six distinct and separable enzymes, but in higher animals reactions 1, 2 and 3 are catalyzed by a trifunctional enzyme, and reactions 5 and 6 by a bifunctional enzyme. What advantages might there be in associations of these active sites?

For dihydroorotate synthetase, the product of reaction 1, carbamoyl phosphate (CAP) is very unstable but is rapidly transformed by aspartate transcarbamoylase which is 50 times more active (per active site) than carbamoyl phosphate synthetase. High levels of carbamoyl aspartate (CA-asp) may be toxic, but this intermediate is rapidly consumed by the high dihydroorotase activity. Because the first three reactions are catalyzed by a single protein, the three enzyme active sites are expressed in a constant ratio under all conditions of growth; this maintains CAP and CA-asp at low levels. For UMP synthase, OMP decarboxylase is far more active (per active site) than orotate PRTase, resulting in low cellular levels of the intermediate, OMP, which would otherwise be subject to enzymatic hydrolysis (in cells from higher animals).

Biosynthesis of Purine Nucleotides

The synthesis of the purine ring is considerably more complex than pyrimidine synthesis. Starting with P-Rib-PP, inosine monophosphate (IMP) is formed in 10 steps (Fig. 15-16). The overall reaction is

$$\text{P-Rib-PP} + 2\text{Gln} + \text{Gly} + \text{2,10-formyl THF} + \text{HCO}_3^- + \text{Asp} + 4\text{ATP}$$
$$\rightarrow \text{IMP} + 2\text{Glu} + 2\text{THF} + \text{fumarate} + 4\text{ADP} + \text{PP}_i$$

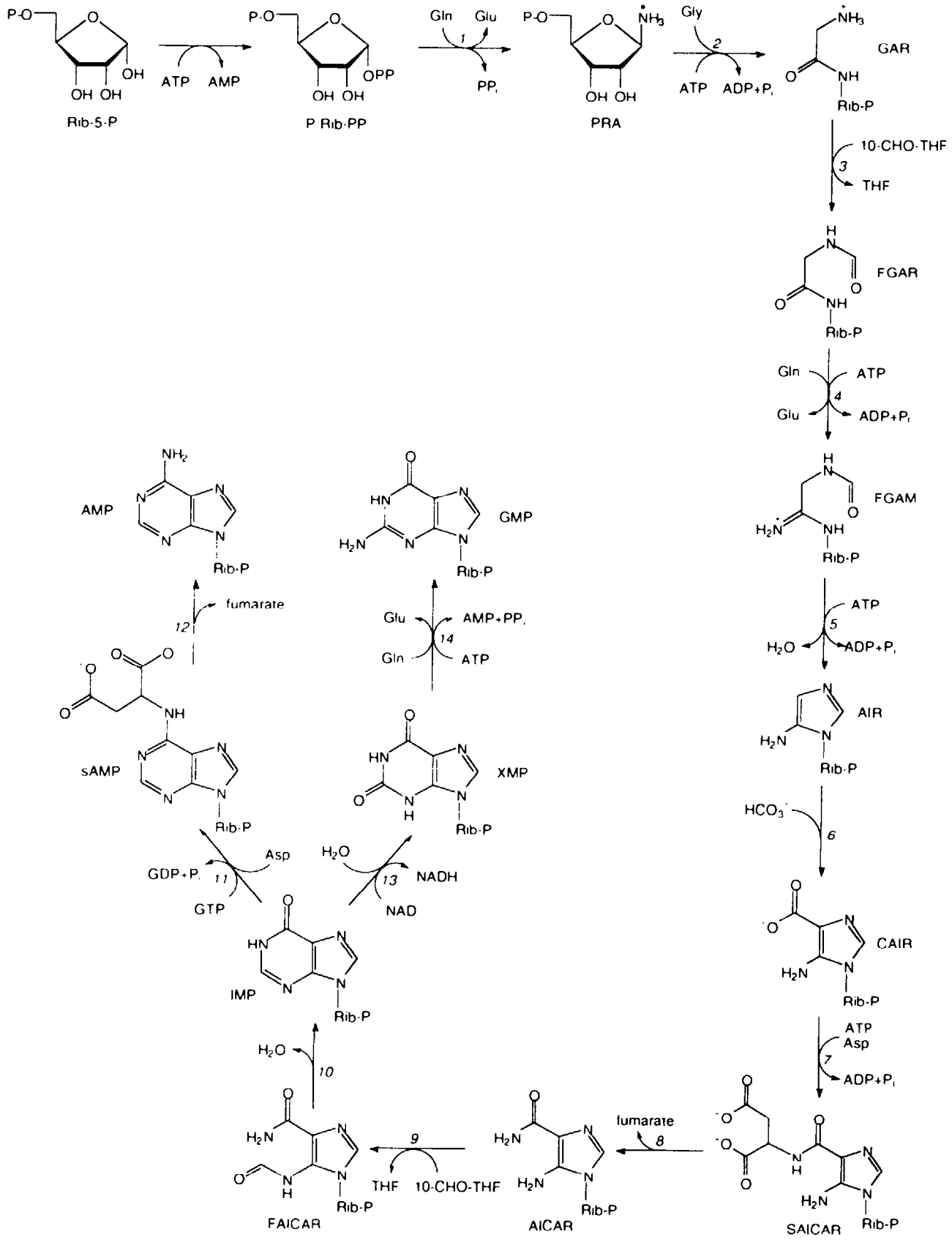
The details of the pathway for *de novo* biosynthesis are shown in Fig. 15-16. The amino acid L-glutamine is the substrate providing nitrogen atoms for reactions 1, 4 and 14, catalyzed by amido phosphoribosyltransferase, FGAM synthetase and GMP synthetase, respectively. Glycine is a substrate at reaction 2, and L-aspartate at reactions 7 and 11. P-Rib-PP is a substrate and activator for amidophosphoribosyltransferase, which is subject to inhibition by AMP, IMP and GMP and by polyglutamate derivatives of dihydrofolate.

The enzymatic activity of amido phosphoribosyltransferase (P-Rib-PP → PRA) is low and flux through the *de novo* pathway *in vivo* is regulated by the end-products, AMP, IMP and GMP. Inhibition of reaction 1 by dihydrofolate polyglutamates would signal the unavailability of N^{10}-formyl tetrahydrofolate, required as a substrate at reactions 3 and 9 of the pathway. The purine pathway is subject to further regulation at the branch point from IMP; XMP is a potent inhibitor of *IMP cyclohydrolase* (FAICAR → IMP), AMP inhibits *adenylosuccinate synthetase* (IMP → sAMP) and GMP inhibits *IMP dehydrogenase* (IMP → XMP).

There are four multifunctional enzymes in the pathway: A *trifunctional enzyme* comprising *GAR synthetase*, *GAR transformylase* and *AIR synthetase* catalyzes reactions 2, 3 and 5 (PRA → GAR → FGAR, FGAM → AIR; Fig. 15-16), respectively. The GAR synthetase and GAR transformylase domains may be separated by digestion of the trifunctional enzyme with the protease, chymotrypsin.

A *bifunctional* enzyme, comprising the activities of *AIR carboxylase* and *SAICAR synthetase*, catalyzes reactions 6 and 7 of the purine pathway (AIR → CAIR → SAICAR; Fig. 15-16). A second bifunctional enzyme, *IMP synthase*, containing the activities of *AICAR transformylase* and *IMP cyclohydrolase*, catalyzes reactions 9 and 10 of the pathway (AICAR → FAICAR → IMP; Fig. 15-16). Human IMP synthase has a subunit molecular weight of 62.1 kDa and associates as a dimer. A

Fig. 15-16 The *de novo* purine biosynthetic pathway. Rib-5-P, ribose 5-phosphate; P-Rib-PP, 5-phosphoribosyl 1-pyrophosphate; PRA, 5-phosphoribosylamine; 10-CHO-FH_4, N^{10}-formyl tetrahydrofolate; GAR, glycineamide ribotide; FGAR, *N*-formylglycineamide ribotide; FGAM, *N*-formylglycineamidine ribotide; AIR, 5-aminoimidazole ribotide; CAIR, 4-carboxy-5-aminoimidazole ribotide; SAICAR, *N*-succino-5-aminoimidazole-4-carboxamide ribotide; AICAR, 5-aminoimidazole-4-carboxamide ribotide; FAICAR, 5-formamidoimidazole-4-carboxamide ribotide; sAMP, *N*-succino-AMP. Enzymes: (1) amido phosphoribosyltransferase; (2) GAR synthetase; (3) GAR transformylase; (4) FGAM synthetase; (5) AIR synthetase; (6) AIR carboxylase; (7) SAICAR synthetase; (8) adenylosuccinase; (9) AICAR transformylase; (10) IMP cyclohydrolase; (11) sAMP synthetase; (12) adenylosuccinase; (13) IMP dehydrogenase; (14) GMP synthetase.


Rib-5-P
ATP
AMP
P Rib-PP
Gln
Glu
1
PP$_i$
PRA
Gly
2
ATP
ADP+P$_i$
GAR
10-CHO-THF
3
THF
FGAR
Gln
Glu
4
ATP
ADP+P$_i$
FGAM
ATP
5
H$_2$O
ADP+P$_i$
AIR
HCO$_3$
6
CAIR
ATP
Asp
7
ADP+P$_i$
SAICAR
fumarate
8
AICAR
9
THF
10-CHO-THF
FAICAR
H$_2$O
10
IMP
GDP+P$_i$
11
GTP
Asp
H$_2$O
13
NADH
NAD
XMP
GMP
Glu
14
Gln
AMP+PP$_i$
ATP
fumarate
12
sAMP
AMP
Rib-P

trifunctional enzyme, *C_1-THF synthase*, containing N^5,N^{10}-methylene tetrahydrofolate (5,10-CH_2-THF) dehydrogenase, N^5,N^{10}-methenyl tetrahydrofolate (5,10,CH-THF) cyclohydrolase and N^{10}-formyl tetrahydrofolate (10-CHO-THF) synthetase, catalyzes the reactions 5,10-CH_2-THF → 5, 10-CH-THF → 10-CHO-THF and THF → 10-CHO-THF. The N^{10}-formyl tetrahydrofolate produced is a substrate for GAR and AICAR transformylases catalyzing reactions 3 and 9 of the pathway (Fig. 15-16). In higher eukaryotes, the dehydrogenase and cyclohydrolase activities are found in one domain of the protein, which is fused to a larger synthetase domain, forming a trifunctional enzyme.

There is a fifth bifunctional enzyme which catalyzes reactions 8 and 12 of the purine pathway (Fig. 15-16) but adenylosuccinate lyase has *one* active site with dual specificity, catalyzing both reactions (SAICAR → AICAR, sAMP → AMP; Fig. 15-16). All 14 enzymatic activities of Fig. 15-16 are cytosolic and there is a variety of evidence for association of subsets of these activities *in vivo*. The existence of a "pathway particle" or "*metabolon*" for *de novo* purine biosynthesis in intact cells has been proposed.

EXAMPLE 15.12

There is some evidence for the existence of a "metabolon" for *de novo* purine biosynthesis which contains the 14 enzymes of the pathway (Fig. 15-16) and four additional enzymes involved in the synthesis of N^{10}-formyl tetrahydrofolate.

Possible selective advantages of this association of catalytic sites in the course of evolution could be:

(*a*) *channeling* of unstable intermediates, such as phosphoribosylamine (PRA) between successive enzymes of the pathway before they diffuse away from the confines of the metabolon;

(*b*) *coordinate regulation* of several enzymatic activities of the metabolon by effectors which bind at a single regulatory site;

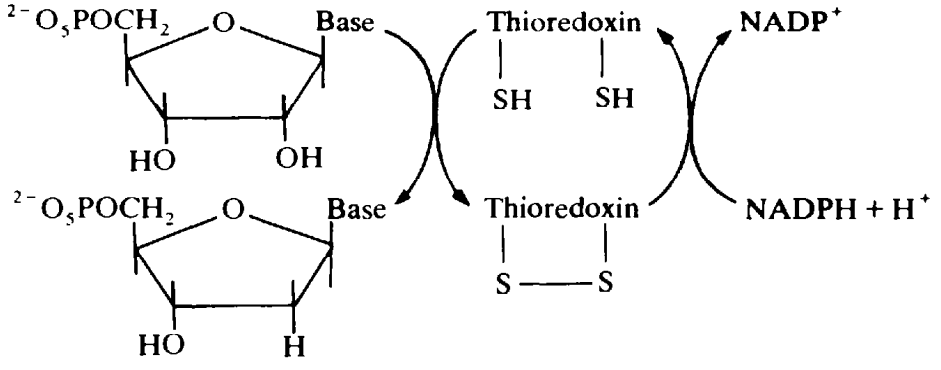
(*c*) *coordinate expression* of enzymatic activities expressed in a single protein, maintaining their catalytic activities in a constant ratio under all conditions of growth (cf. dihydroorotate synthetase).

Deoxyribonucleotide Synthesis

The synthesis of DNA is dependent on a ready supply of deoxyribonucleotides. The substrates for these are the ribonucleoside diphosphates ADP, GDP, CDP, and UDP; the enzyme responsible for the reduction of these substrates to their corresponding deoxy derivatives is *ribonucleotide reductase*, which has *thioredoxin* as a cosubstrate.

EXAMPLE 15.13

Thioredoxin is a protein of 12,000 Da that can donate two electrons by the oxidation of two cysteine sulfhydryl groups to cystine; oxidized thioredoxin is reduced by NADPH:



The overall reaction for the synthesis of, for example, deoxyadenosine diphosphate (dADP) is:

$$ADP + NADPH + H^+ \rightarrow dADP + NADP^+ + H_2O$$

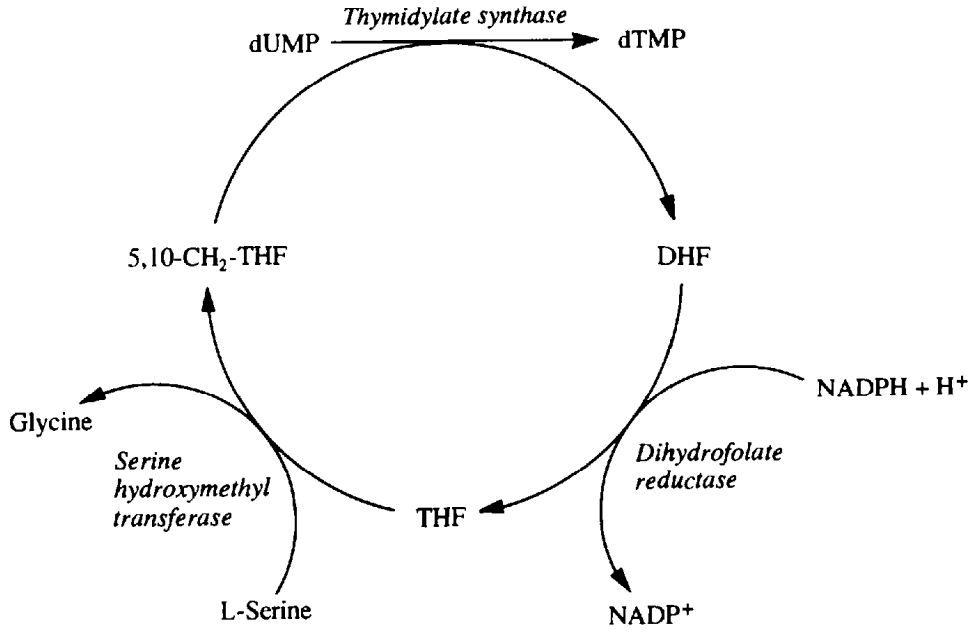
The deoxyribonucleoside diphosphates are phosphorylated by ATP.

EXAMPLE 15.14

How is the activity of ribonucleotide reductase regulated to ensure that the deoxyribonucleotides are produced in the correct ratio?

The enzyme contains two catalytic sites, two regulatory sites and two specificity sites. The catalytic site binds the substrates, thioredoxin (reduced by $NADPH + H^+$) and the nucleoside diphosphates. The allosteric regulatory site binds ATP as an *activator* in competition with dATP as an *inhibitor*. The specificity site binds dGTP, dTTP and dATP but not dCTP and modulates ribonucleotide reductase activity selectively for the four NDP substrates to balance the four dNTP pools.

Cells making DNA must also be able to make deoxythymidine triphosphate (dTTP). The key step in the synthesis of dTTP is the conversion of dUMP to dTMP via *thymidylate synthase*. The reaction requires a source of *N^5,N^{10}-methylene tetrahydrofolate* (see Sec. 15.7, Fig. 15-19) to provide the methyl group. In this reaction, the tetrahydrofolate is oxidized to dihydrofolate. Dihydrofolate must be reduced to tetrahydrofolate via the enzyme *dihydrofolate reductase* so that more N^5,N^{10}-methylene tetrahydrofolate can be made from serine in a reaction catalyzed by *serine hydroxymethyltransferase*. These three reactions, which are essential for the formation of dTMP, are shown below.



Nucleotide Antagonists as Anticancer Drugs

To grow and divide, cancer cells must duplicate their chromosomes which are composed of deoxynucleoside 5′-monophosphates (dNMPs) polymerized in unique sequences. Cancers may differ from normal cells of the body by growing more rapidly and/or by "cycling" and dividing continuously. Inhibitors of nucleotide biosynthesis will have *selective toxicity* for such cancers due to a depletion or imbalance in the cellular levels of dNTPs required for DNA synthesis, which would be more pronounced than in normal cells. A selective depletion of one of the four dNTPs (e.g., dTTP) by treatment of cancer

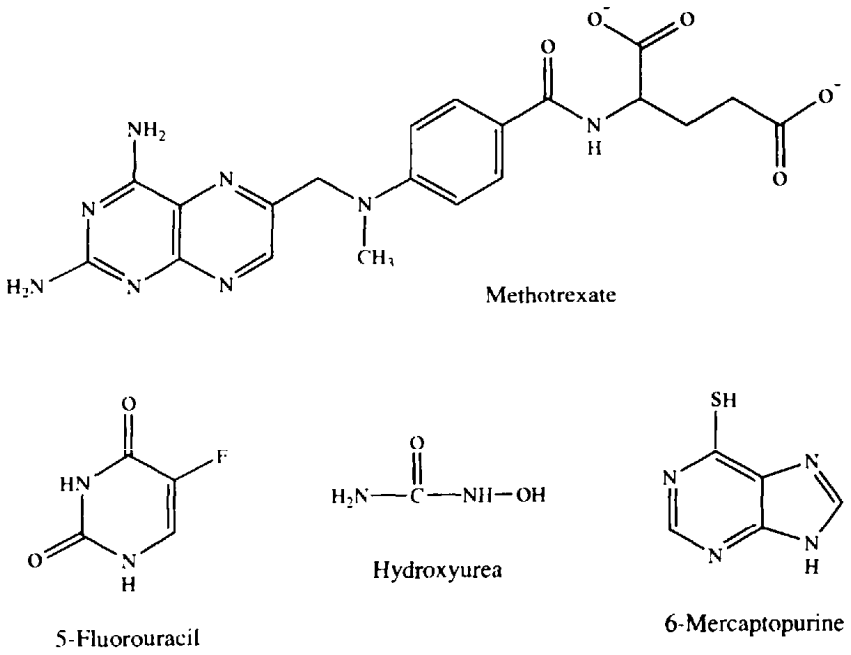


Fig. 15-17 Chemical structures of 4 anticancer drugs which inhibit nucleotide biosynthesis.

cells with a drug (e.g., 5-fluorouracil) may lead to the arrest of DNA synthesis in the cancer cell and consequent cell death. Alternatively, if dTTP is not absent, but decreased, the imbalance in the cellular pools of dNTPs may lead to genetic miscoding with consequent fatal mutations.

The chemical structures of four commonly used anticancer drugs are shown in Fig. 15-17. Methotrexate was the first "true" anticancer drug, synthesized in 1949, and has been in clinical use for treatment of a variety of cancers since the early 1950s. Methotrexate is a potent inhibitor of *dihydrofolate reductase* with an inhibition constant (K_I) for interaction with the enzyme of $10^{-9}\,M$. Inhibition of this enzyme in a cell leads to major accumulation of DHF to concentrations of $\sim 2.5\,\mu M$, and minor decreases in THF. Marked decreases in THF may not be seen due to the release of bound THF in methotrexate-treated cells. The high levels of DHF are toxic to the cell, inhibiting the reaction catalyzed by thymidylate synthase,

$$\text{dUMP} + 5,10\text{-CH}_2\text{-THF} \rightarrow \text{dTMP} + \text{DHF}$$

and the first reaction of *de novo* purine biosynthesis catalyzed by amido PRTase,

$$\text{P-Rib-PP} + \text{L-glutamine} \rightarrow \text{PRA} + \text{L-glutamate} + \text{PP}_i$$

In leukemia cells treated with methotrexate, levels of dTTP decrease, and there may be less marked decreases in dATP and dGTP resulting from inhibition of amido PRTase. The consequent imbalance in nucleotide pools results in genetic miscoding and cell death.

EXAMPLE 15.15

Methotrexate remains an anticancer drug of major importance for "combination chemotherapy." A number of mechanisms by which cancer cells gain resistance to methotrexate have been identified.

(*a*) Amplification of the gene which encodes the target enzyme, dihydrofolate reductase.

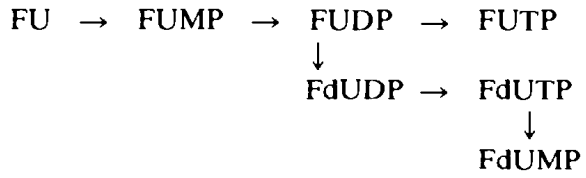
(*b*) Mutation of the folate transporter which translocates methotrexate into cells.

(*c*) Mutation of dihydrofolate reductase so that binding of the substrate, dihydrofolate, is retained but the binding of methotrexate is weaker.

(*d*) Loss of activity of the enzyme folylpolyglutamyl synthetase; this enzyme adds a polyglutamyl tail to methotrexate, thus retaining it in the cancer cell.

Such methotrexate-resistant cells are found in cancer patients who had been given methotrexate as a single agent, where a remission is followed by relapse with drug-resistant cancer.

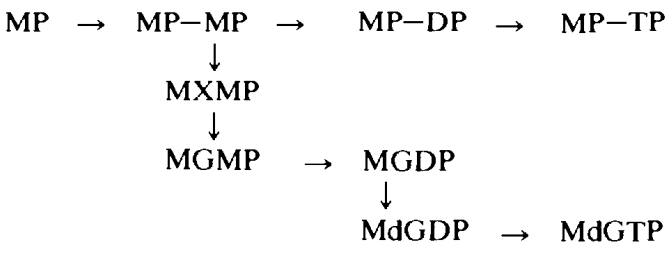
5-Fluorouracil (FU) is also a very useful anticancer drug which is taken up by cells and metabolized in the following ways.



The primary mechanism of action of FU may be inhibition of thymidylate synthase (dUMP → dTMP) by 5-fluorodeoxy UMP (FdUMP). The FdUMP binds to thymidylate synthase with the other substrate 5,10-CH_2-THF forming a tight *ternary complex*. A cysteine residue at the active site of the enzyme attacks the natural substrate, dUMP, forming a transient covalent bond. With FdUMP, this bond cannot be broken due to the presence of the 5-fluoro group on the pyrimidine ring (Fig. 15-17). Thymidylate synthase is thus permanently inactivated by this *suicide inhibitor* with a consequent depletion of dTMP and thus dTTP in cells. However, 5-fluorouracil may also kill cancer cells by two other mechanisms. FUTP which accumulates in cells (see above) may be incorporated into RNA causing genetic miscoding, or FdUTP may be incorporated into DNA, again leading to fatal mutations.

Hydroxyurea is a simple molecule (Fig. 15-17) which inhibits ribonucleotide reductase. This enzyme accepts the four NDPs, UDP, CDP, ADP and GDP, as substrates and reduces them to the corresponding dNDP. The mechanism of catalysis involves the formation of an unusual *tyrosyl radical cation* which then induces formation of a radical form of the NDP substrate. Hydroxyurea quenches this tyrosyl radical cation intermediate leading to depletion of all four dNTPs required as substrates for the synthesis of DNA.

6-Mercaptopurine (MP; Fig. 15-17) is one of many drugs discovered by the Nobel Laureates Gertrude Elion and George Hitchings. 6-Mercaptopurine was synthesized in the early 1950s and remains a useful anticancer drug. Like 5-fluorouracil, 6-mercaptopurine may have several mechanisms of toxicity which may depend upon the type of cell. 6-Mercaptopurine may enter cells and be metabolized in the following way:



The 6-mercaptopurine 5′-monophosphate (MP-MP) formed is a potent inhibitor of amido PRTase, and thus *de novo* purine biosynthesis (Fig. 15-16) is blocked. The 6-mercaptodeoxy GTP (MdGTP) is incorporated into DNA and causes genetic miscoding.

Salvage Synthesis of Nucleotides

The *de novo* synthesis of pyrimidines and purines, particularly purines, is energetically expensive and hence most (80%) of the purines and pyrimidines obtained from the degradation of nucleic acids,

particularly RNA, are salvaged for reuse. Human cells contain three *phosphoribosyltransferases* (PRTases) which convert preformed nucleobases to the equivalent nucleoside 5′-monophosphate (NMP). They are adenine-, hypoxanthine-guanine- and orotate- (uracil) PRTases:

$$\text{Nucleobase} + \text{P-Rib-PP} \rightarrow \text{NMP} + \text{PP}_i$$

P-Rib-PP is an activated form of ribose 5-phosphate (Rib-5-P). The equilibria for these reactions favor nucleobase + P-Rib-PP, but the pyrophosphate (PP_i) formed is hydrolyzed to phosphate by pyrophosphatase, and NMP is therefore formed from the corresponding nucleobase.

While the PRTases salvage nucleobases within cells, nucleosides such as adenosine and uridine are present in the blood at much higher concentrations ($\sim 1\ \mu M$) than the equivalent nucleobases, adenine and uracil. Indeed, the brain synthesizes pyrimidine nucleotides (UTP and CTP) via salvage synthesis from uridine produced by the liver and released into the circulation. Human cells may contain at least three types of nonspecific nucleoside transporters, and nucleosides are internalized more rapidly than nucleobases.

Once inside the cell, the nucleoside is converted to the corresponding NMP, adenosine by adenosine kinase and uridine by uridine kinase:

$$\text{Nucleoside} + \text{ATP} \rightarrow \text{NMP} + \text{ADP}$$

The NMP salvaged by a PRTase or kinase reaction may then be converted to the nucleoside 5′-triphosphate (NTP):

$$\text{NMP} \rightarrow \text{NDP} \rightarrow \text{NTP}$$

EXAMPLE 15.16

In some diseases, excessive amounts of purines are produced in the body, leading to accumulation of urate. Patients with *Lesch-Nyhan syndrome* lack the enzyme hypoxanthine-guanine phosphoribosyltransferase (HG-PRTase). Children born with this disorder are mentally retarded and prone to self-mutilation. They produce excessive amounts of purines due to accumulation of P-Rib-PP which stimulates the first enzyme of the pathway, amido PRTase (Fig. 15-16). The excess purines are degraded via the reactions

$$\text{Adenosine} \rightarrow \text{Inosine} \rightarrow \text{Hypoxanthine} \rightarrow \text{Xanthine} \rightarrow \text{Urate}$$

Patients with Lesch-Nyhan syndrome may also suffer from gout, which is due to an accumulation of urate in the body with deposition of crystals of sodium urate in the joints and kidneys. However, gout more commonly occurs due to failure of urate excretion by the kidneys, or due to accumulation of P-Rib-PP for reasons other than a deficiency of HG-PRTase.

Nucleic Acid Degradation

The general scheme for the degradation of nucleic acids has much in common with that of proteins. Nucleotides are produced by hydrolysis of both dietary and endogenous nucleic acids. The *endogenous* (cellular) polynucleotides are broken down in lysosomes. DNA is not normally turned over rapidly, except after cell death and during DNA repair. RNA is turned over in much the same way as protein. The enzymes involved are the *nucleases*; *deoxyribonucleases* and *ribonucleases* hydrolyze DNA and RNA, respectively, to oligonucleotides which can be further hydrolyzed (Fig. 15-18), so eventually purines and pyrimidines are formed.

Most of the enzymes involved in the hydrolysis of dietary DNA and RNA are secreted in the intestine. Thus, ribonucleases are found in the lumen of the small intestine, while the phosphodiesterases and nucleotidases are present in the mucosal cells.

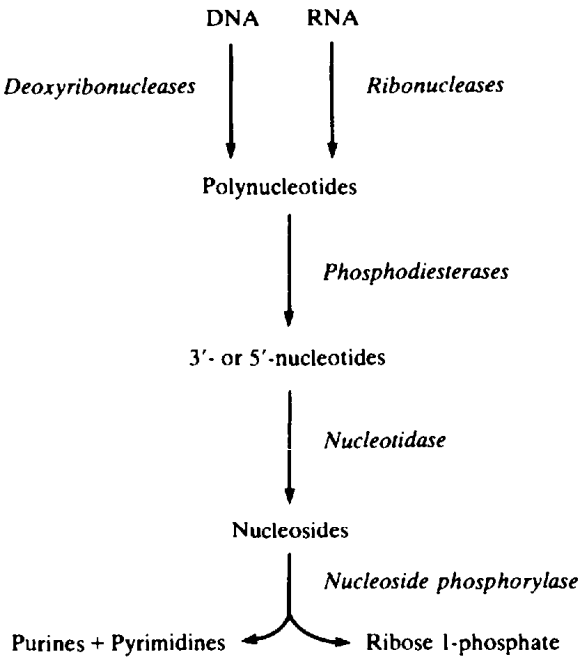
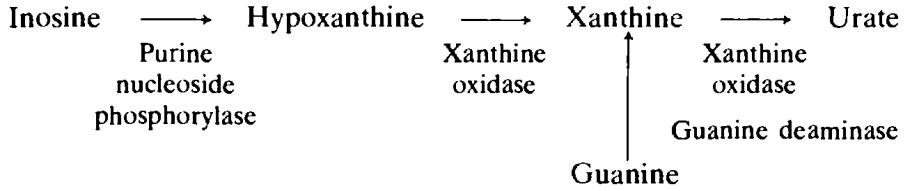


Fig. 15-18 Hydrolysis of DNA and RNA to purines and pyrimidines.

Degradation of Purines and Pyrimidines

Purines and pyrimidines in excess of cellular requirements can be degraded. The extent of degradation depends on the organism. Humans cannot degrade purines beyond uric acid because we lack the enzyme *uricase*, which splits the purine ring to form *allantoin*. In humans excess AMP is deaminated to IMP by the action of a specific deaminase. IMP is then hydrolyzed by 5'-nucleotidase to form inosine. Inosine and guanine are oxidized to urate as follows:



Thus, because we lack uricase, we excrete, albeit in small amounts, uric acid every day. The liver synthesizes about 0.8 g of uric acid per day, but 20–50 percent enters the gut in the gastric juice and in the bile and is degraded by microorganisms. For some animals (the *uricoteles*, such as birds), uric acid is the form in which excess nitrogen is excreted. Unless there are enzymes missing, as is the case of humans, nonuricotelic organisms can degrade purines to urea, ammonia, and carbon dioxide. Pyrimidines are also degraded to urea and ammonia.

15.7 METABOLISM OF C_1 COMPOUNDS

Folic Acid Derivatives

Several processes described above use one-carbon derivatives of *tetrahydrofolate* (Fig. 15-19). For example, the synthesis of the purine ring (Fig. 15-16) requires *N^{10}-formyl tetrahydrofolate*. Thymidylate synthase, a key enzyme in pyrimidine synthesis, uses *N^5-N^{10}-methylene tetrahydrofolate* both as a substrate and as a reducing agent. This compound, perhaps the most important in C_1 metabolism, is

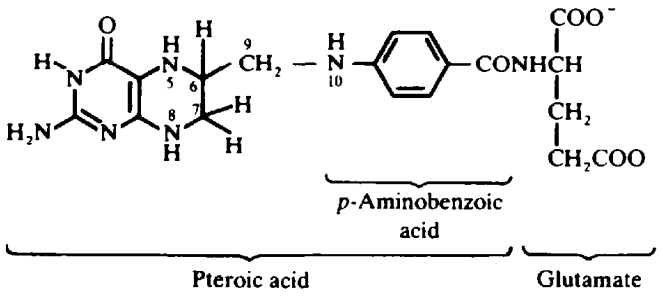
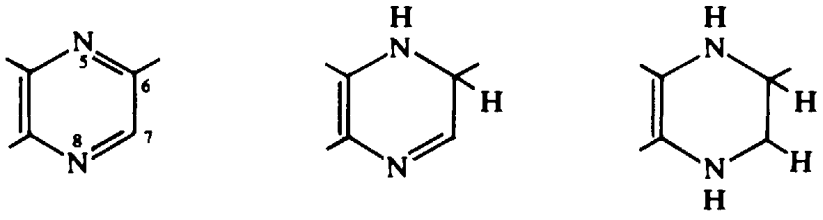


Fig. 15-19 Tetrahydrofolic acid.

also involved in the interconversion of serine and glycine. All these compounds are derivatives of 5,6,7,8-tetrahydrofolic acid, which is the reduced form of the vitamin *folate* (or folic acid).

The pteroic acid moiety of tetrahydrofolate consists of a reduced *pteridine ring* and *p*-aminobenzoic acid. Folate from the diet is absorbed by the intestinal mucosa and in two enzymatic steps is reduced to tetrahydrofolate which is the active form of the coenzyme. Mammals cannot synthesize folate; this normally does not present a problem because microorganisms of the intestinal tract readily do so.

The two steps in the reduction of folic acid to tetrahydrofolate are catalyzed by *dihydrofolate reductase*. Both of these reactions require NADPH as a source of electrons.



Folate **Dihydrofolate** **Tetrahydrofolate**

In order to understand the bewildering variety of reactions involving tetrahydrofolate, it is essential to realize that in biological systems, one-carbon compounds may exist in *five* different oxidation states. The most *reduced* form is methane, CH_4, and the most *oxidized* form is CO_2. In between these two extremes are *methanol* (CH_3OH), *formaldehyde* (CH_2O), and *formate* ($HCOO^-$).

Question: Are all the above one-carbon compounds involved in "C_1 metabolism"?

Methane and carbon dioxide are the exceptions. Methane is an end product of anaerobic metabolism of many microorganisms, and carbon dioxide (for carboxylation) is handled by biotin-containing enzymes.

Table 15.4 lists the various one-carbon groups carried by tetrahydrofolate derivatives.

These C_1 groups may be attached to N atoms in positions 5 or 10 (Fig. 15-19) or may form a bridge between the two. N^5-methyl-THF is formed in mammals by a virtually irreversible reaction

Table 15.4. Tetrahydrofolate (THF) Derivatives

Group Carried	THF Derivative
$-CH_3$	N^5-Methyl-THF
$-CH_2OH$	N^5,N^{10}-Methylene-THF
$-CHO$	N-Formyl-THF
$-CH=$	N^5,N^{10}-Methenyl-THF

that is catalyzed by the enzyme *methylene-THF reductase*; the other THF derivatives are interconverted through a series of oxidation-reduction and hydration-dehydration reactions:

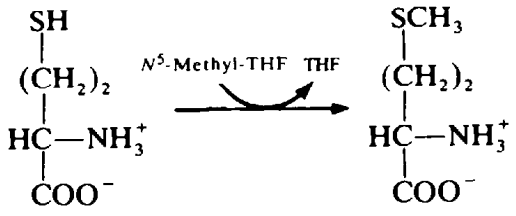
$$N^5\text{-Methyl-THF} \leftarrow N^5,N^{10}\text{-Methylene-THF} \rightleftharpoons N^5\text{-}N^{10}\text{-Methenyl-THF} \rightarrow N^{10}\text{-Formyl-THF}$$

Question: What is the major reaction that replenishes C_1 units in THF?

With the exception of N^5-methyl-THF, the THF derivatives are directly synthesized from a C_1 unit in the appropriate oxidation state, and THF. The major *anaplerotic* reaction is that catalyzed by *serine hydroxymethyltransferase* (Sec. 15.1).

Biological Methylations

The transfer of methyl groups is a common biochemical reaction, and the introduction of a methyl group into a molecule is an important way of modifying biological activity, as in the case of epinephrine versus norepinephrine. The methyl groups originate from N^5-methyltetrahydrofolate, although this compound is involved directly in only one methylation reaction. The simplest form of this reaction occurs in plants, where it is catalyzed by the enzyme *homocysteine transmethylase*:



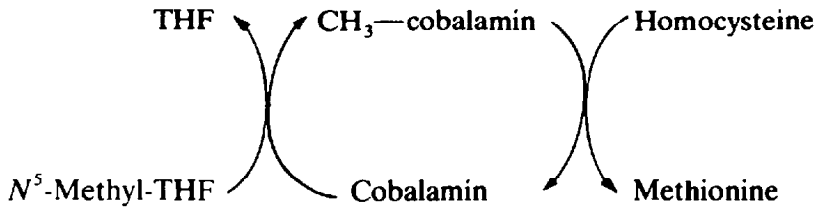
Homocysteine Methionine

The mammalian synthesis of methionine is more complex and requires cobalamin, a coenzyme form of vitamin B_{12}. Note that because methionine is an essential amino acid, it must be supplied in the diet; methionine that is used for methylation (Fig. 15-20) is degraded to homocysteine, and this is remethylated to give methionine. These reactions merely *recycle* methionine and do *not* constitute a means of *net* synthesis.

EXAMPLE 15.17

Vitamin B_{12} does not exist in plants, and strict vegetarians risk suffering from vitamin B_{12} deficiency. Thus, we are dependent on animal or bacterial sources for our vitamin B_{12}.

Cobalamin is a complex molecule containing a Co atom. In the mamalian synthesis of methionine, cobalamin acts as a coenzyme by accepting the methyl group from N^5-methyltetrahydrofolate and transferring it to homocysteine. The reaction is catalyzed by *cobalamin–N^5-methyl-THF: homocysteine methyltransferase*. The overall reaction is



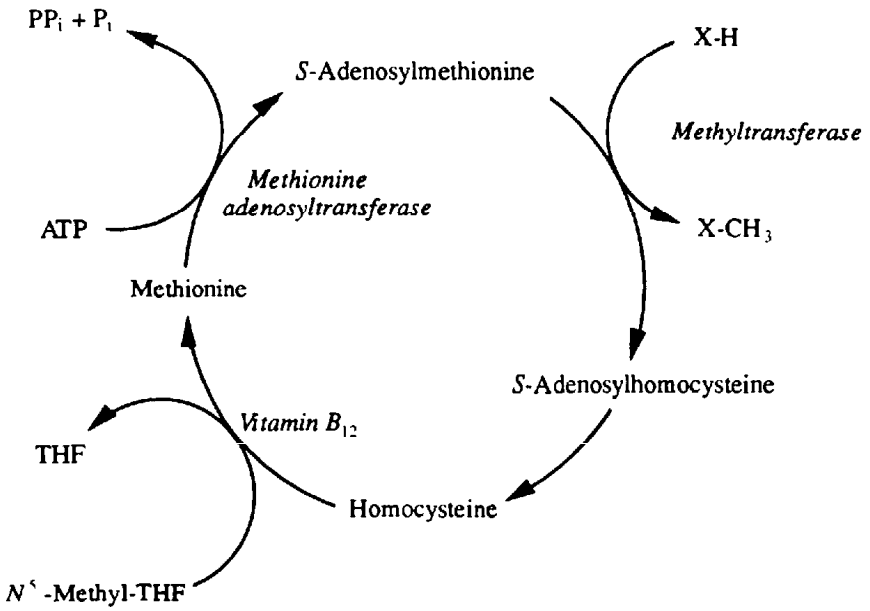


Fig. 15-20 Methionine metabolism.

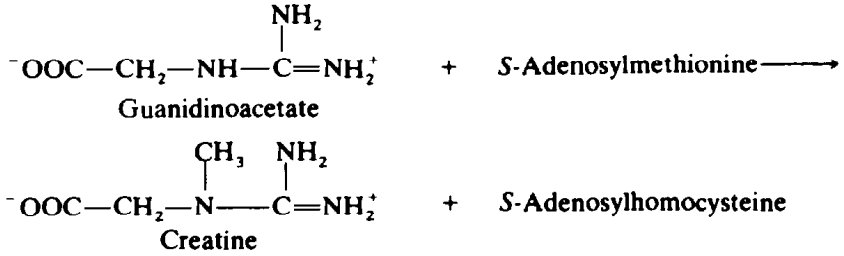
EXAMPLE 15.18

One aspect of vitamin B_{12} deficiency is that it results in the accumulation of N^5-methyl-THF. N^5-Methyl-THF is synthesized in mammals by an irreversible reaction (as shown above); if it cannot be utilized because of a deficiency of vitamin B_{12}, then it accumulates. This causes a depletion of the other forms of THF, resulting in a deficiency of THF. *Megaloblastic anemia* (pernicious anemia) is associated with a deficiency of cobalamin and THF.

The methyl group on methionine is activated when methionine is converted to S-*adenosylmethionine*. It is the methyl group of *S*-adenosylmethionine that is the immediate donor in biological methylations. Important reactions in which *S*-adenosylmethionine acts as the methyl donor are the synthesis of creatine, epinephrine, and phosphatidylcholine.

EXAMPLE 15.19

Creatine is synthesized from guanidinoacetate (which is made from glycine and arginine).



The carbon skeleton of homocysteine is salvaged and used to synthesize methionine (Fig. 15-20). Alternatively, homocysteine is used to synthesize cysteine.

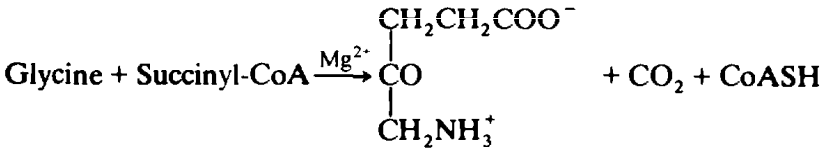
The importance of the cycle depicted in Fig. 15-20 is that it conserves homocysteine. Methionine and cysteine that are used for protein synthesis will deplete homocysteine from the cycle so at least methionine must be replaced in the diet.

15.8 PORPHYRIN METABOLISM

The synthesis and turnover of porphyrins, heme precursors, are important because of the central roles of the heme proteins, hemoglobin, and the cytochromes. Quantitatively, hemoglobin synthesis is a major part of the nitrogen economy in humans.

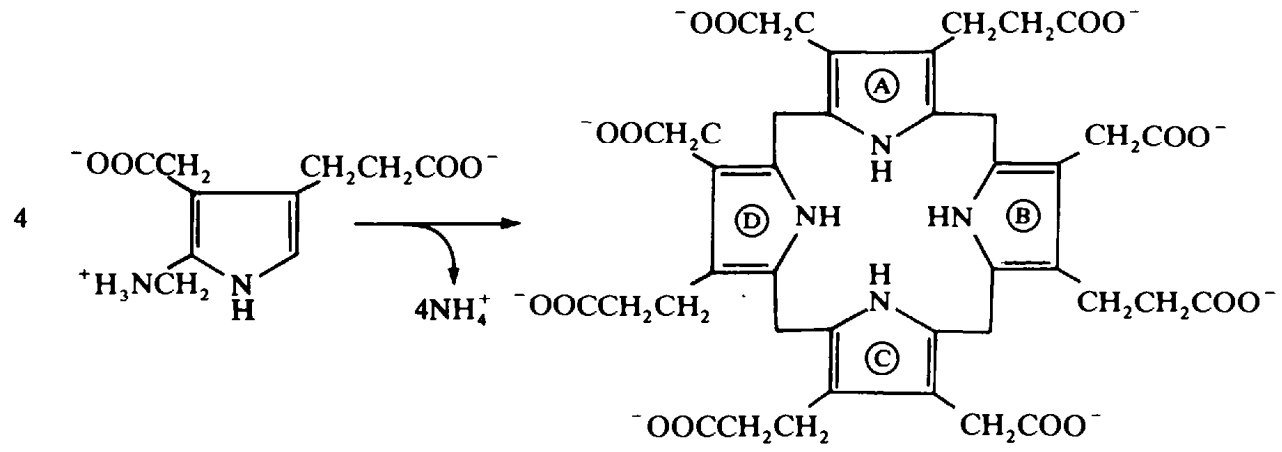
Porphyrin Synthesis

The first step of porphyrin synthesis is the condensation of succinyl-CoA and glycine to form *δ-aminolevulinate*. The reaction takes place in mitochondria, where succinyl-CoA is available. The reaction is irreversible and requires pyridoxal phosphate and Mg^{2+}. It is catalyzed by the enzyme *δ-aminolevulinate synthase*.



δ-Aminolevulinate

Subsequent reactions occur in the cytoplasm and they are irreversible. Two molecules of δ-aminolevulinate are condensed by the enzyme *porphobilinogen synthase* to form the trisubstituted pyrrole *porphobilinogen*. Two enzymes, *uroporphyrinogen synthase* and *uroporphyrinogen cosynthase*, condense four molecules of porphobilinogen to the porphyrin *uroporphyrinogen III*.



Porphobilinogen Uroporphyrinogen III

Note that uroporphyrinogen III is not a symmetrical molecule. During its synthesis, one of the pyrrole rings (ring D) is reversed, with the result that the acetate and propionate side chains are not symmetrically arranged around the porphyrin ring. The key porphyrin intermediate in cytochrome and hemoglobin synthesis is *protoporphyrin IX* (Fig. 15-21).

The synthesis of protoporphyrin IX involves two alterations to the side chains of uroporphyrinogen III: decarboxylation of the acetate groups to methyl groups and decarboxylation of the propionate residues in rings A and B to vinyl groups (—CH=CH_2). The first decarboxylations take place in the cytoplasm, while the formation of the vinyl groups and the conversion of the methylene bridges (—CH_2—) to the unsaturated methene (=CH—) is mitochondrial. The final product from these reactions is the fully aromatic, planar protoporphyrin IX. The final reaction in the mitochondria is the chelation of Fe^{2+} to form heme, a reaction which occurs spontaneously, although the enzyme *ferrochelatase* enhances the rate.

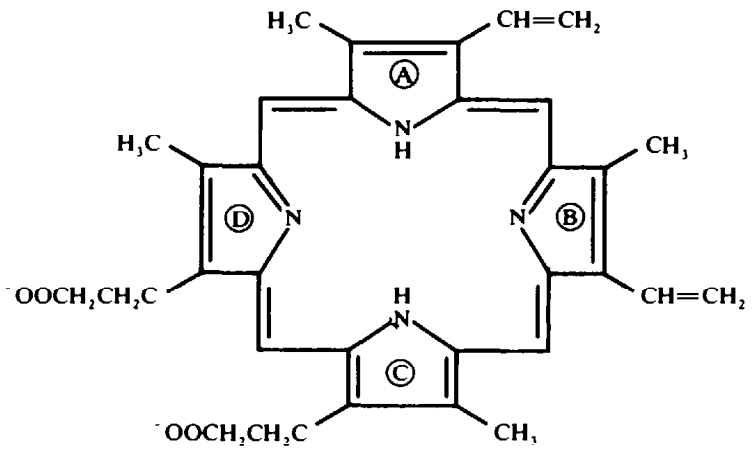


Fig. 15-21 Structure of protoporphyrin IX.

Heme is the functional group in hemoglobin and myoglobin, the cytochromes, and the enzymes catalase and peroxidase. These molecules all have quite different functions: hemoglobin carries oxygen; myoglobin stores oxygen; the cytochromes transfer electrons; and catalase and peroxidase are enzymes that catalyze the decomposition of hydrogen peroxide and oxidation by peroxides, respectively.

EXAMPLE 15.20

What determines the function of the heme group in different proteins?

Although there are differences in the way the heme is attached to the proteins, it is the amino acid sequences of the proteins that determine the function of the porphyrins in these molecules.

Regulation of Heme Synthesis in Reticulocytes

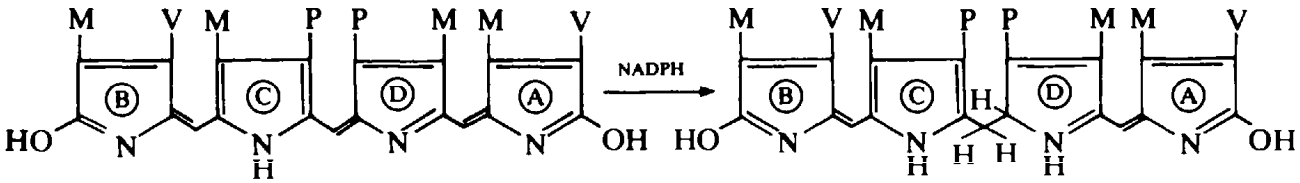
Heme synthesis is controlled primarily by *δ-aminolevulinate synthase* (ALA synthase). There are two mechanisms of control, and each involves a process that affects the *concentration* of the enzyme. First, the half-life of ALA synthase, as shown by experiments in rat liver, is very short (60–70 min). Like many mitochondrial proteins, ALA synthase is encoded by nuclear genes, synthesized on cytoplasmic ribosomes, and the enzyme is translocated into the mitochondria. The second and main regulating factor is the inhibition of ALA synthase by *hemin*. Hemin differs from heme in that the Fe atom is in the Fe^{3+} oxidation state. Heme spontaneously oxidizes to hemin when there is no globin to form hemoglobin. Hemin serves a second function in the regulation of hemoglobin synthesis in reticulocytes. It controls the synthesis of globin.

High concentrations of hemin inhibit the transport of ALA synthase into the mitochondria, where one of the substrates, succinyl-CoA, is formed. Thus, heme synthesis is inhibited until enough globin is made to react with any heme already formed. Low concentrations, or the absence, of hemin is the signal that globin is not needed; this protein (and, therefore, globin) synthesis is inhibited. In the absence of hemin, a *protein kinase* is activated; this phosphorylates an *initiation factor* of (eukaryotic) protein synthesis, *eIF-2*, which then inhibits polypeptide chain initiation (Chap. 17) and hence inhibits globin synthesis.

Degradation of Hemoglobin

The life span of the human erythrocyte is ~120 days with about 0.85 percent of the total being broken down each day in the reticuloendothelial (RE) cells of the spleen, liver, and bone marrow.

The erythrocytes are lysed inside the digestive vesicles of the reticuloendothelial cells and the hemoglobin is degraded. The globin is hydrolyzed to amino acids, and the heme is metabolized as follows. The porphyrin ring is oxidatively cleaved between rings A and B to form the linear tetrapyrrole *biliverdin* (which is a green color). The complete reaction requires molecular oxygen and NADPH, and the final product is bilirubin (which is an orange-red color). The Fe^{2+} is salvaged via transferrin and stored within the protein *apoferritin*, and the methene bridge between rings A and B is removed as CO.



Biliverdin Bilirubin

In the structures shown, M = methyl group, V = vinyl group, and P = propionate. *Biliverdin* and *bilirubin* are bile pigments and are familiar as the green and orange coloring of bruises. Bilirubin, a water-insoluble molecule, is released into the plasma complexed with albumin and transported to the liver. In the liver it is solubilized by being converted to *bilirubin diglucuronide* (90 percent) and *bilirubin sulfate* (10 percent).

EXAMPLE 15.21

Glucuronic acid is a derivative of glucose. The $-CH_2OH$ group on C-6 has been oxidized to $-COOH$. The active form of glucuronic acid is UDP-glucuronate and this is used in creating glucuronides, such as bilirubin diglucuronide. Glucuronide formation (and sulfate formation) is a common means of increasing solubility because of the polarity of the $-OH$ and $-COO^-$ groups of glucuronic acid. It is particularly important in the excretion of many insoluble drugs via the bile.

Bilirubin diglucuronide is excreted from the liver via the bile into the intestine. Within the bowel, it is hydrolyzed, and the bilirubin is reduced to *urobilinogen* and *stercobilinogen*. These are excreted in urine as *urobilin* after reabsorption from the bowel, and in feces as *stercobilin*; these pigments give urine and feces their characteristic colors.

Solved Problems

SYNTHESIS AND DIETARY SOURCES OF AMINO ACIDS

15.1. What effect will high concentrations of NADPH and 2-oxoglutarate have on the assimilation of ammonia?

SOLUTION

Either or both of these compounds in high concentration will favor glutamate synthesis in the reaction catalyzed by glutamate dehydrogenase. This shift of the chemical reaction to the right will result in assimilation of ammonia.

$$NH_4^+ + \text{2-Oxoglutarate} + NADPH \rightleftharpoons \text{Glutamate} + NADP^+ + H_2O$$

15.2. How does the TCA cycle operate if oxaloacetate and 2-oxoglutarate are removed for amino acid synthesis?

SOLUTION

Any oxaloacetate or 2-oxoglutarate removed from the TCA cycle must be replaced. Pyruvate is converted by *pyruvate carboxylase* to oxaloacetate, which can then enter the TCA cycle to yield 2-oxoglutarate.

15.3. In the fed state, what is the direction of the net carbon flux in transamination?

SOLUTION

In the fed state, when there is abundant protein and carbohydrate, dietary protein is hydrolyzed to amino acids. Those amino acids not required for protein synthesis are converted to 2-oxoacids by the aminotransferases. The 2-oxoacids are then converted into lipids and carbohydrate for storage. *Glutamate dehydrogenase* catalyzes the formation of ammonia from the excess amino groups derived from the amino acids; this ammonia is excreted as urea.

15.4. What determines the fate of pyruvate produced in muscle from amino acid metabolism?

SOLUTION

In the normal fed state, pyruvate is oxidized via the pyruvate dehydrogenase complex (PDH), but in starvation, PDH is inactivated; thus, pyruvate is converted into alanine (Fig. 15-16) which enters the blood and is conveyed to the liver, where gluconeogenesis takes place.

DIGESTION OF PROTEINS

15.5. What is the fate of the proteases after protein digestion?

SOLUTION

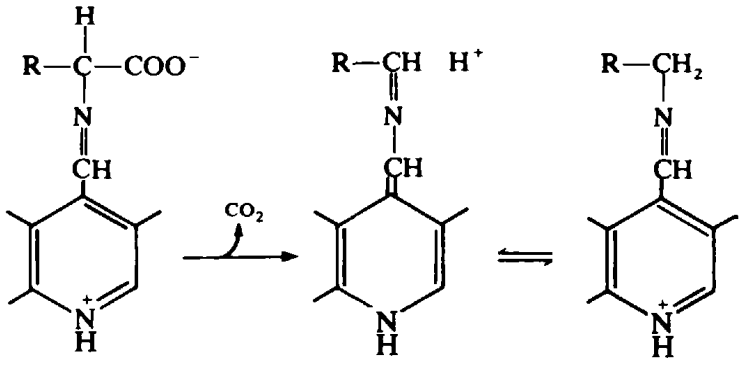
The proteases and other pancreatic enzymes such as pancreatic lipase are eventually degraded (by proteases). Digestion is responsible for the turnover of about 50 g of endogenous protein per day. This comes from the breakdown of the pancreatic enzymes as well as the epithelial cells of the gut, which are replaced every 24 hours.

AMINO ACID CATABOLISM

15.6. Pyridoxal phosphate is a coenzyme in amino acid decarboxylations. What is a likely mechanism of the decarboxylation, and what are the products?

SOLUTION

As with transaminations, the first step is the formation of a Schiff base between the amino acid and pyridoxal phosphate:



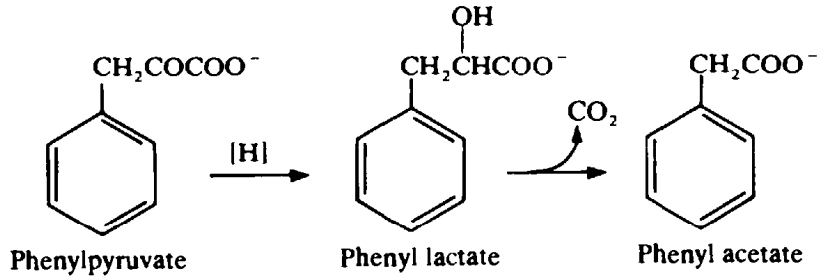
The quaternary nitrogen acts as an electron sink, which facilitates the decarboxylation. Further electron and proton shifts produce a Schiff base between the amine and pyridoxal phosphate, which is then hydrolyzed.

Amino acid decarboxylations are involved in the synthesis of several metabolically important amines, e.g., *5-hydroxytryptamine* (serotonin) from tryptophan, *histamine* from histidine, and *γ-aminobutyric acid* (GABA) from glutamate.

15.7. The disease phenylketonuria, which causes severe mental retardation, is characterized by the urinary excretion of phenylpyruvate. Why is this formed?

SOLUTION

Phenylketonuria is due to an inborn error of phenylalanine metabolism. Typically, it is due to a deficiency of phenylalanine hydroxylase. Atypically, it can be caused by a deficiency of *dihydrobiopterin reductase* and a resultant inability to synthesize *biopterin*. All these conditions cause an accumulation of phenylalanine, which can be transaminated to phenylpyruvic acid.



Phenylpyruvate can be reduced to phenyl lactate and oxidatively decarboxylated to phenylacetate, both of which are also excreted in the urine.

DISPOSAL OF EXCESS NITROGEN

15.8. Under what conditions would citrulline, argininosuccinate, or arginine accumulate in the liver?

SOLUTION

The accumulation of any of these amino acids could be due to reduced activity of their respective enzymes in the urea cycle (Sec. 15.5), resulting in decreased overall activity of the cycle. Inborn errors of metabolism are known for deficiencies in these enzymes. Decreased activity of the urea cycle results in elevated levels of ammonia in the blood, a condition known as *hyperammonemia* that causes nausea, vomiting and even coma.

15.9. Can hyperammonemia be controlled?

SOLUTION

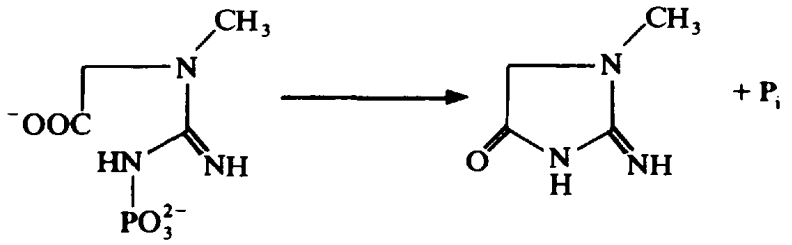
Feeding low-nitrogen diets that contain the 2-oxoacid counterparts of the essential amino acids will reduce ammonia concentrations.

15.10. Predict the form in which excess ammonia would be excreted in the following organisms: tadpoles, frogs, birds, and mammals.

SOLUTION

As aquatic animals, tadpoles can use diffusion and dilution to remove excess ammonia, but the amphibian frog excretes urea. Birds and some reptiles excrete uric acid. In most mammals, excess ammonia is excreted as urea.

15.11. Creatinine is an important nitrogenous constituent of urine. It is formed nonenzymatically from phosphocreatine primarily in muscle by spontaneous cyclization, resulting from phosphorolysis.



Phosphocreatine **Creatinine**

Is the formation of creatinine significant?

SOLUTION

The rate of formation of creatinine is constant; the amount is dependent on muscle mass. Hence, an individual excretes a constant amount (1–1.5 g per day for an adult man). Its excretion represents a net loss of methyl groups.

PURINE AND PYRIMIDINE METABOLISM

15.12. Uric acid crystals can be deposited in the joints, causing the painful condition known as *gout*. What, in turn, causes the deposition of the uric acid crystals?

SOLUTION

Gout is caused either by an overproduction of purines, which leads to the overproduction of uric acid, or to a failure of the kidneys to excrete uric acid. Because of its insolubility, uric acid precipitates in the joints and causes inflammation.

15.13. The *Lesch-Nyhan syndrome* is a distressing disorder that includes neurological abnormalities, self-mutilation, and overproduction of uric acid. Why is uric acid overproduced?

SOLUTION

The Lesch-Nyhan syndrome is caused by a deficiency of the phosphoribosyltransferase that is involved in the salvage pathway for hypoxanthine and guanine (HGPRTase). The accumulation of P-Rib-PP stimulates purine biosynthesis.

15.14. What causes hereditary orotic aciduria?

SOLUTION

A genetic deficiency of the orotate phosphoribosyltransferase results in accumulation of orotate and its excretion in the urine.

15.15. Why does a deficiency of adenosine deaminase result in severe combined immunodeficiency?

SOLUTION

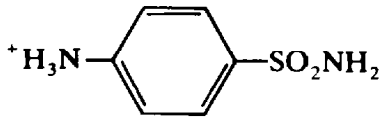
Patients that lack adenosine deaminase are unable to degrade adenosine to inosine, or deoxyadenosine to deoxyinosine. Accumulated deoxyadenosine (dAdo) is converted to nucleotides as follows:

$$\text{dAdo} \rightarrow \text{dAMP} \rightarrow \text{dADP} \rightarrow \text{dATP}$$

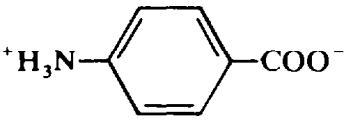
Millimolar levels of dATP kill T- and B-lymphocytes, resulting in immunodeficiency.

METABOLISM OF C_1 COMPOUNDS

15.16. The sulfonamides are antibacterial agents which are S analogs of *p*-aminobenzoate. The simplest of these is sulfanilamide.



Sulfanilamide



p-Aminobenozate

How do the sulfonamides work?

SOLUTION

They act as competitive inhibitors of the incorporation of *p*-aminobenzoate into folic acid by bacteria. Without folic acid, the bacteria cannot grow.

PORPHYRIN METABOLISM

15.17. Inborn errors of metabolism are normally associated with a deficiency or absence of an enzyme. Can excessive production of an enzyme also result in an inborn error of metabolism?

SOLUTION

Excessive production of liver δ-aminolevulinate synthase causes two forms of congenital *porphyria*. These diseases are characterized by overproduction of porphyrins and excretion of large amounts of δ-aminolevulinate and porphobilinogen. Some ethnic groups have a high incidence of this disease, and in these people, acute attacks are brought on by barbiturates and other compounds that induce synthesis of the enzyme.

Supplementary Problems

15.18. Why would a diet rich in energy but low in nitrogen result in manutrition?

15.19. Would you expect a diet in which meat had been replaced by cheese to be nutritionally adequate in nitrogen?

15.20. Which proteases would be needed to hydrolyze the following peptides completely?

(*a*) Tyr-Phe-Gly-Ala

(*b*) Ala-Arg-Tyr-Glu

(*c*) Leu-Trp-Lys-Ser

15.21. Glutamine from the diet and from muscle during fasting can be converted into alanine in the intestine. Which enzymes are involved in this conversion?

15.22. Why is arginine classified as an essential amino acid when it is synthesized in the urea cycle?

Chapter 16

Replication and Maintenance of the Genetic Material

16.1 INTRODUCTION

DNA is the *genetic material* of cells; i.e., it contains the *genes*. An individual gene is represented by a segment of DNA (up to several kb), that is eventually *expressed* as a polypeptide, frequently a structural protein or enzyme. The DNA is present in the chromosomes of a cell, with one or more chromosomes comprising the *genome*. The genome represents a single complement of the genetic information of a particular cell. Each chromosome contains a large number of genes. During cell division, the chromosomal DNA must produce exact replicas (copies) of itself for segregation and *partitioning*, into daughter cells. This production of copies of the DNA is known as *replication* and involves the synthesis of new DNA chains.

For normal cell growth and proliferation, the DNA must be protected from various types of damage. Such damage, induced, for example, by uv irradiation, can involve the chemical alteration of the DNA and, consequently, deleterious *mutation*. Cells are able to correct or *repair* such damage. One of the best-understood mechanisms of repair involves the synthesis of new DNA, which replaces the damaged portion. This is called *repair synthesis* of DNA. The extent of repair synthesis is very small in comparison with the DNA synthesis accompanying replication of the chromosomes.

Most cells (bacteria, plants, and animals) contain one, two, or a small number of most genes. However, in response to certain situations, some cells will produce many copies of a particular gene. This is known as *gene* or *DNA amplification* and involves several-fold replication of specific segments of the DNA.

EXAMPLE 16.1

Faults in DNA replication and repair in animal cells can lead to a predisposition to cancer. This helps to emphasize the importance of these processes in biology. However, much of our understanding of how they occur at the molecular level has depended upon studies with the simpler bacteria.

16.2 SEMICONSERVATIVE REPLICATION OF DNA

DNA in the chromosomes of most organisms is *double-helical*; i.e., it consists of two polydeoxynucleotide chains (or strands) twisted around one another in the form of a helix. The genetic information is contained in the *sequence* of nucleotides along one of the chains, with the sequence in one being complementary to that in the other (Chap. 7). A *replica* of DNA is one that is an exact copy of itself.

Question: What feature of the DNA structure provides the basis for reproduction of the original nucleotide sequence in a replica?

Each chain of double-helical DNA is bound to the other through complementary base pairs, with adenine (A) in one being hydrogen-bonded to thymine (T) in the other, and guanine (G) to cytosine (C). Watson and Crick proposed that, to achieve precise copying of a nucleotide (base) sequence, the two chains of the DNA must *unwind* from one another to allow each single chain to act as a *template* for the synthesis of a new one. Thus, the assembly of the sequence in the newly synthesized

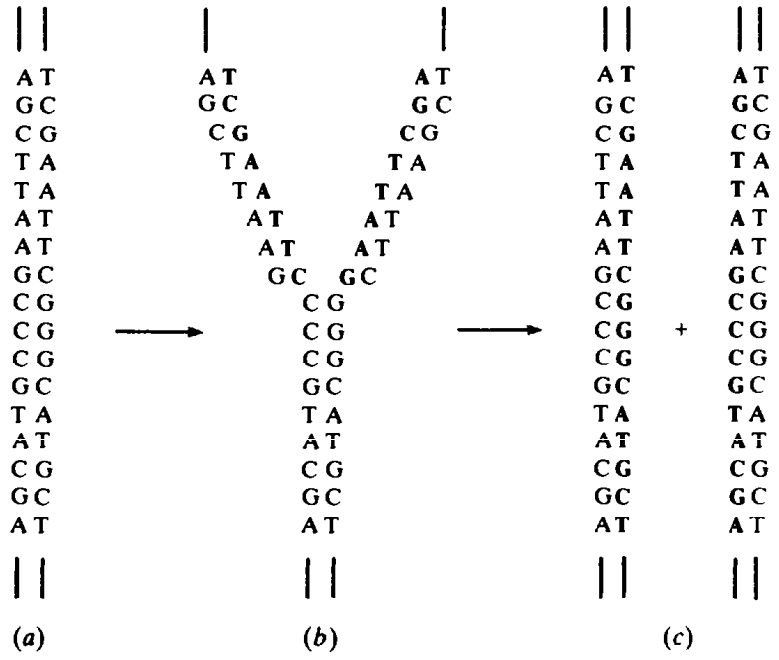


Fig. 16-1 Unwound DNA as a template for replication. The bold letters indicate newly synthesized DNA.

chain is determined by the base-pairing specificity of the sequence in the template. This is illustrated in Fig. 16-1 (the DNA *duplex* is shown as a ladder-type structure).

In Fig. 16-1, the DNA to be replicated (*a*) is shown to unwind from the top, and the partially replicated structure (*b*) appears Y-shaped and contains a *replication fork*. Within each arm of this structure, the newly synthesized chains (shown in bold type) are assembled according to the base-pairing instructions of the unwound template chains. The length of sequence shown in (*a*) would represent only a very minor segment of a giant DNA molecule in a chromosome. The completely replicated segments are shown in (*c*); they are now separated from each other. Each contains one of the original chains and one new chain. The *parental* segment of DNA gives rise to two *daughter* segments, which will remain intact through subsequent generations.

Question: What is meant by the term *semiconservative replication* of DNA?

Semiconservative replication refers to the conservation of just one half of the parental DNA structure when it undergoes replication to give two daughter molecules. Thus, in Fig. 16-1, each chain of the parental DNA acts as a template and remains intact through the *doubling* process.

16.3 TOPOLOGY OF DNA REPLICATION

A chromosome contains a single DNA molecule, which is generally very large; e.g., some bacterial chromosomes are composed of as many as 4×10^6 base pairs. Furthermore, in many cases, the DNA is a closed or *circular* structure. Some bacterial chromosomes are linear. We will concentrate first on the topology of replication of the circular bacterial (*E. coli*) chromosome. Little is known about linear bacterial chromosomes in this regard.

Question: What is the form taken by a replicating bacterial chromosome (circular DNA molecule)?

The replicating bacterial chromosome remains in a closed form, with a portion of its length duplicated and joined to the rest of the DNA at replication *forks* (Fig. 16-2).

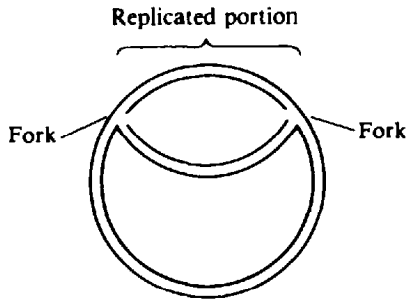


Fig. 16-2 Replicating circular bacterial chromosome.

Question: The replicating bacterial chromosome shown in Fig. 16-2 contains *two* forks. Are both of these *replication* forks?

Yes, both are actively involved in DNA replication and move at approximately equal rates in opposite directions around the circular molecule. This is known as *bidirectional replication.* The replicated portion of the molecule is referred to as a *replication bubble* or *eye form* (because of its appearance in diagrams). The size of the bubble varies from being extremely small up to nearly twice the size of the nonreplicating chromosome. Obviously, the site on the circular molecule of a very small bubble represents the region within which replication was initiated.

Question: Does initiation of replication of the bacterial chromosome occur at a fixed or a variable position?

Genetic studies have established that initiation of replication occurs at a fixed site, called the *initiation site* or *origin of the chromosome* (*oriC*). The nucleotide sequence in this region binds to various proteins to initiate the two forks.

Question: What happens when the two replication forks meet one another at the position opposite *oriC*?

The approaching forks meet, thus releasing two completed circular molecules (daughters). This is called *terminaton of replication.* The forks meet and fuse in what is called the *terminus region* (*tr*) and there is a specific mechanism to ensure that the fusion site is restricted to this region.

EXAMPLE 16.2

Initiation of replication at a fixed site on the chromosome and the movement of replication forks away from this site toward the terminus region, implies a fixed order of replication of genes during this process. This has been confirmed through measurements of the relative *copy number* of genes in DNA isolated from growing cells. Genes near the origin of replication occur at a higher average frequency than those near the terminus.

Replication of the bacterial chromosome can be divided into three stages: *initiation*, *elongation*, and *termination* (Fig. 16-3). *Initiation* refers to the generation of replication forks at the origin. *Elongation* describes the progression of these forks around the chromosome, with concomitant DNA synthesis or *chain growth. Termination* refers to the fusion of the approaching forks, which results in two completed chromosomes that can separate from one another.

The term *replicon* is used to describe a *unit of replication* that is under the control of a single origin. Thus, the bacterial chromosome is a replicon; its replication is under the control of *oriC*. This will be treated in more detail below.

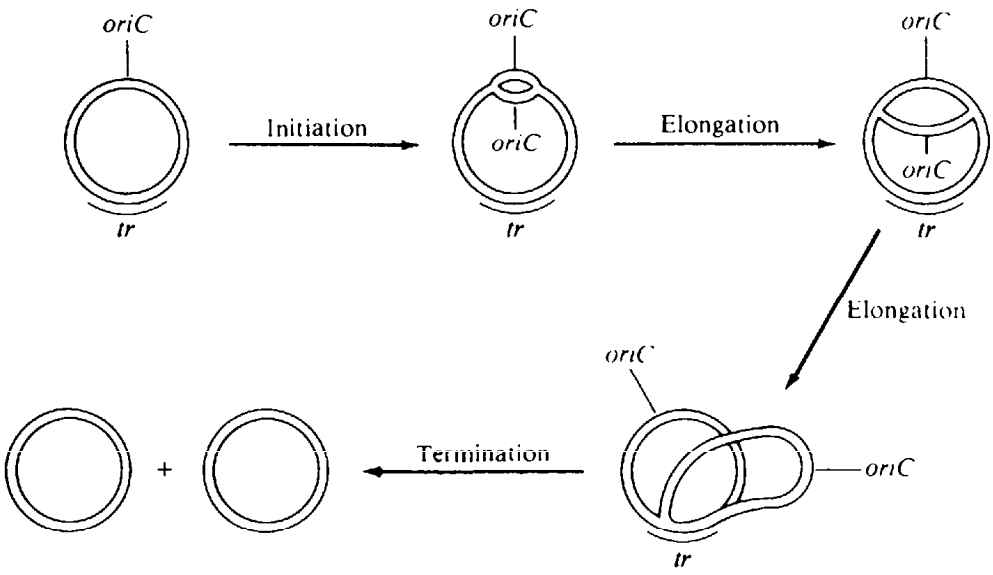


Fig. 16-3 Stages in the replication of a bacterial chromosome.

Question: Do all circular, double-stranded DNA molecules replicate in the same manner as the bacterial chromosome?

Not exactly, in that there is a defined origin of replication, yes. However, in some cases, e.g., some *plasmids*, replication proceeds in a single unique direction around the DNA molecule. This is called *unidirectional replication*. In other situations, e.g., mitochondrial DNA (Chap. 1), there is a more complex sequence of events which involves the unidirectional copying of only one strand over most of the DNA length, in a different type of eye form, referred to as a *D loop*.

Replication of the ~5,000-kb bacterial chromosome takes about 40 min and occurs *throughout* the bacterial-division cycle. Thus, *each* fork replicates about 50 kb of DNA per minute. In eukaryotic cells, DNA replication is restricted to the portion of the mitotic cell-division cycle called the *S phase*, which can extend for several hours.

EXAMPLE 16.3

The cell cycle in eukaryotes is divided into four phases, G_1, S, G_2, and M (Fig. 16-4). The cycle starts at the beginning of G_1 (G = gap), which often constitutes the major phase in duration. DNA replication commences at the beginning of the S phase and is completed before entry into G_2. Mitosis, or the M phase, is relatively short and includes those steps that lead to chromosome segregation and partitioning into two daughter cells.

Question: A eukaryotic cell contains ~1,000 times the amount of DNA of a bacterial chromosome. How can replication of eukaryotic DNA be accomplished within the period of the S phase (a few hours)?

Examination of replicating eukaryotic DNA by electron microscopy shows the presence of many *tandemly* arranged bubbles separated by only 30–300 kb and clustered in groups of 20–80 in various regions of the DNA. Both forks of each bubble represent sites of replication, which move in opposite directions, until they fuse with an approaching fork from an adjacent bubble. Thus, eukaryotic DNA contains many tandemly arranged *replicons*, each with an origin of replication. Actually the rate of fork movement in eukaryotes is less than 10 percent of the rate in bacteria, but the genome can be replicated in the time available because of the multitude or origins.

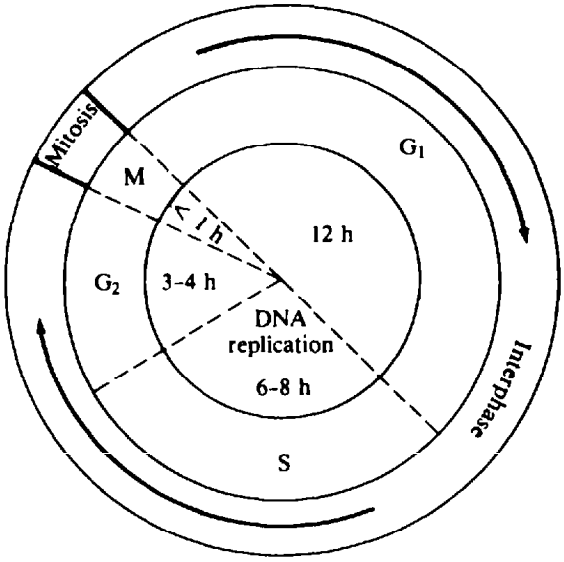


Fig. 16-4 The eukaryotic cell cycle. The durations of the four phases as shown are typical of higher eukaryotic cells growing in tissue culture.

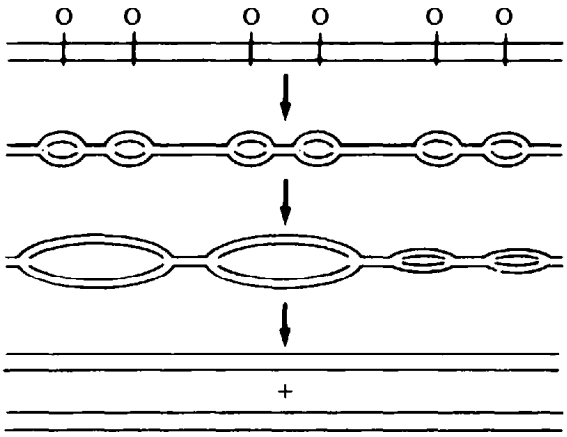


Fig. 16-5 Tandemly arranged replicons in eukaryotic DNA generate a chain of replication bubbles. O denotes an origin of replication.

EXAMPLE 16.4

A cultured mammalian cell contains about 20,000 origins. As shown in Fig. 16-5, the replication forks initiated at each origin, O, have only a relatively short distance to traverse before fork fusion (or termination) with an approaching fork from an adjacent replicon.

16.4 CONTROL OF DNA REPLICATION

The rate of DNA replication is coordinated with the rate of cell division. Thus, a bacterial culture growing in a rich medium has a short generation-time and must accomplish chromosome replication

more quickly than one in a poor medium, where the generation time might be three to four times longer.

Question: By what mechanism is the rate of DNA replication in bacteria altered?

Under all nutrient conditions at a fixed temperature, the rate of replication-fork movement remains fairly constant. The overall rate of DNA replication is determined by the *frequency of initiation* at *oriC*. To accelerate replication, initiation at daughter origins occurs before the ongoing *cycle of replication* is completed. This results in the formation of a *multiforked chromosome*, as shown in Fig. 16-6. Note that a multiforked chromosome contains at least four copies of *oriC*.

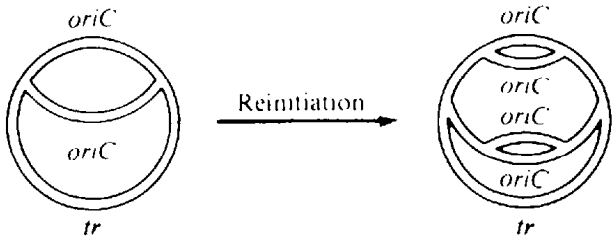


Fig. 16-6 A multiforked chromosome generated by reinitiation of replication at *oriC*.

The control of DNA replication in eukaryotes is more complex than in bacteria. Once a cell enters S phase, the multitude of individual replicons throughout the genome probably replicate in a defined order. It is known that, in some situations, the duration of S phase can vary greatly.

EXAMPLE 16.5

The first two nuclei formed after fertilization of an egg of *Drosophila melanogaster* undergo division in just a few minutes; thus, the length of S phase is *very* short. In this situation, the origins of adjacent replicons are spaced at much shorter than normal intervals, 7–8 kb. The whole genome is replicated quickly because each replication fork traverses only a few kb before fusion with an approaching one.

Question: Can individual replicons in eukaryotes undergo reinitiation, as in bacteria, to yield multiforked structures?

No, but a situation very similar to it can arise in some specialized cells in which specific genes are *amplified*. For example, the developing egg of many animals contains at least a thousand copies of the rRNA genes (Chaps. 7 and 17). In *Drosophila*, the DNA sequences coding for the *chorion* (egg) *proteins* are amplified about 30-fold just before they are needed. To achieve amplification, a small region of the DNA traversing the particular gene (or genes) is replicated many times during a single cell generation, as shown in Fig. 16–7.

Note that the amplified DNA is attached to the chromosomal DNA (Fig. 16-7). When the amplified DNA is *freed* from the chromosome, the freed form is referred to as *extrachromosomal DNA*. Such DNA is generated by *nonhomologous recombination* within an amplified segment to produce a *circular excised sequence*. The sequence in this circular DNA can then be amplified further through a *rolling-circle* mechanism to yield multiple copies of it that are tandemly arranged within a long linear DNA molecule. It is in this form that amplified rRNA genes exist.

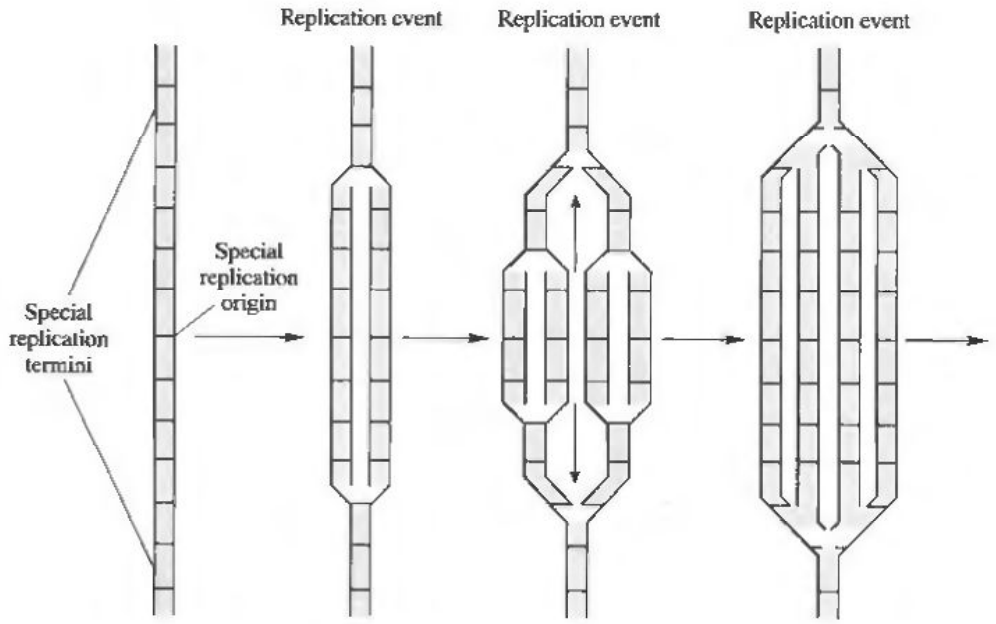


Fig. 16-7 Amplification of a segment of DNA through replication.

EXAMPLE 16.6

The rolling-circle mechanism of DNA replication operates as shown in Fig. 16-8.

A cut or *nick* is made in one of the two chains, and the circle rolls, peeling away one end of the cut chain, to yield the equivalent of two single chains, which function as templates for DNA synthesis at a fork. As the circle continues to roll, a linear duplex molecule containing multiple copies of the sequence is generated. The rolling-circle mechanism is also seen as a stage in the replication of some viral DNA molecules.

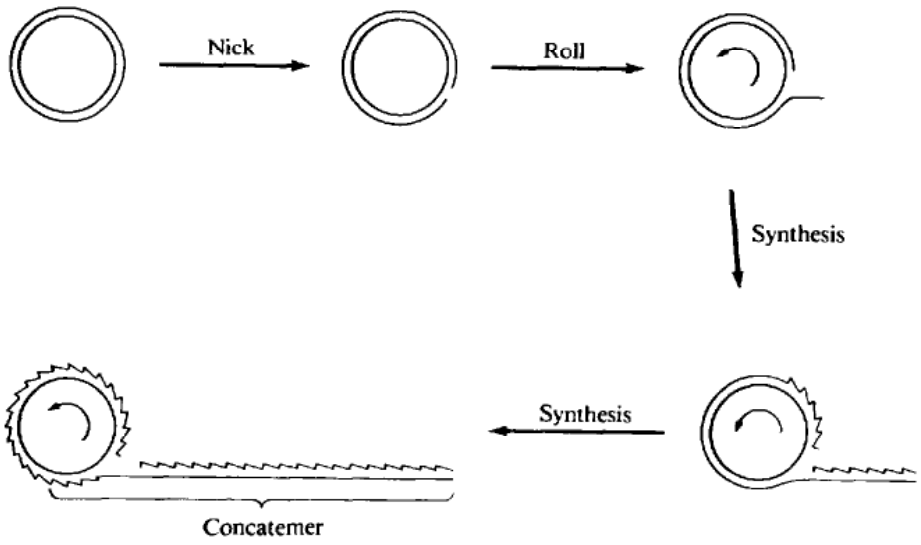


Fig. 16-8 Rolling-circle mechanism of DNA replication.

16.5 ENZYMOLOGY OF DNA REPLICATION IN BACTERIA

As already discussed, the replication of a *replicon* can be divided into three stages: initiation, elongation and termination. During the elongation phase, DNA chain growth takes place at

replication forks. It is a good stage for examining some of the important enzymes and other proteins involved in replication. The process as it occurs in the bacterium *E. coli* is best understood, and serves as the prototype for other systems. Several enzymes and proteins are involved.

Question: Which enzyme is responsible for the synthesis of new DNA chains at the replication fork?

The enzyme, which uses unwound, single-stranded DNA as a *template*, is called a *DNA polymerase*. There are three distinct DNA polymerases in *E. coli*: DNA polymerase I, II, and III. DNA polymerase I is the most abundant, and DNA polymerase III the least abundant. These two enzymes have important roles in the overall process of DNA replication. The role of DNA polymerase II has not yet been clearly established.

DNA Polymerase I

DNA polymerase I of *E. coli* has been studied in considerable detail. It uses the four deoxynucleoside triphosphates (dNTPs) and Mg^{2+} to catalyze the template-directed polymerization of nucleotides (Fig. 16-9). In the reaction, an incoming dTTP is positioned opposite adenine (A) in the template (through base pairing), and a phosphodiester bond is formed through a *nucleophilic attack* by the 3′ hydroxyl of the growing chain on the α phosphorus of the incoming triphosphate. Pyrophosphate (PP_i) is released. The next nucleotide is then incorporated to extend the chain farther, and so on. Growth is exclusively in the 5′→3′ direction. Another important feature is that the enzyme can add only to a *preexisting* chain (*primer*); it cannot start a new one. Rapid hydrolysis of the released PP_i by *pyrophosphatase* in the cell serves to drive the reaction in the direction of DNA chain growth. The overall reaction catalyzed by DNA polymerase I is:

$$(\text{dNMP})_n + \text{dNTP} \overset{Mg^{2+}}{\rightleftharpoons} (\text{dNMP})_{n+1} + PP_i$$

However, this is only one of *three* types of reactions catalyzed by DNA polymerase I. Its two other activities involve the *hydrolysis* of phosphodiester bonds. One is a *5′→3′ exonuclease* acting on double-stranded DNA, and the other is a *3′→5′ exonuclease* acting on a *frayed* or *mismatched* terminus of double-stranded DNA. DNA polymerase I is thus a *multifunctional* enzyme. It consists of a single

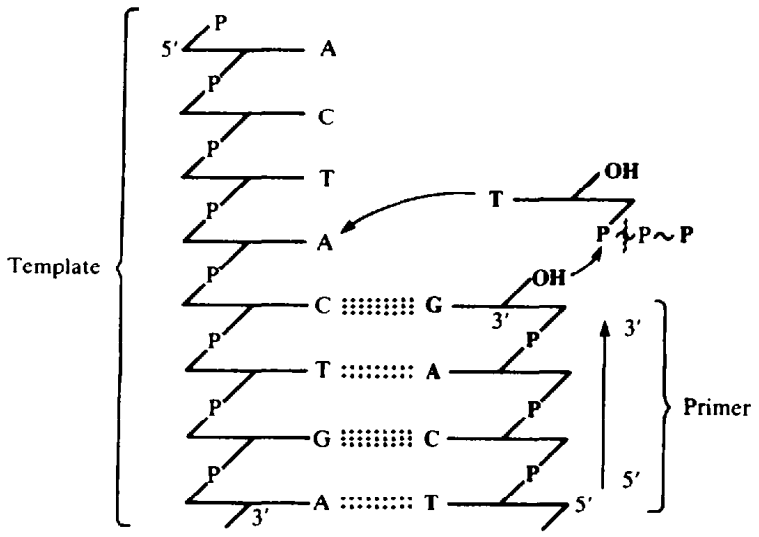


Fig. 16-9 Template-directed addition of a nucleotide unit to a growing DNA chain by DNA polymerase I. The bold letters indicate newly synthesized DNA. The rows of dots denote hydrogen bonds.

polypeptide chain (M_r = 109,000). Its cleavage by proteases yields a *large* fragment (*Klenow fragment*; M_r = 76,000) which contains the polymerase and 3′→5′ exonuclease activities, and a *small* fragment (M_r = 35,000) which contains only the 5′→3′ exonuclease activity.

DNA Polymerase III

DNA polymerase III is also a multifunctional enzyme. It resembles DNA polymerase I in catalytic properties; however, there are slight differences with respect to the type of template primer preferred for DNA synthesis as well as the preferred substrates for the two exonuclease activities. It contains many polypeptide subunits (α, ε, θ, τ, γ, δ, χ, ψ and β). The complex containing all subunits ($M_r \sim 900{,}000$) is called the DNA polymerase III *holoenzyme*, while that comprising just α, ε and θ exhibits the polymerase activity and is referred to as the *core* enzyme. The holoenzyme carries out most of the DNA synthesis at the replication fork *in vivo*.

Recall that the two chains in double-stranded DNA have opposite polarities (Chap. 7); one is 5′→3′ in direction and the other is 3′→5′. But the DNA polymerases can extend a chain only in the 5′→3′ direction, so at a replication fork, only one of the new chains can be made 5′→3′ and move in the direction of fork movement. In Fig. 16-10, this occurs in the upper arm.

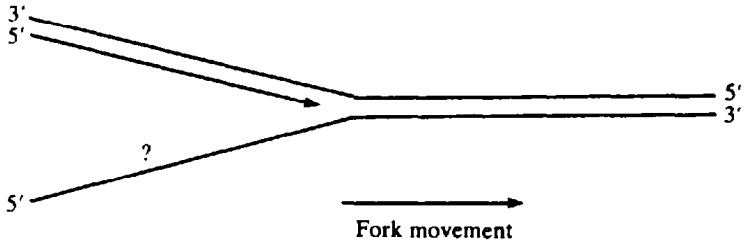


Fig. 16-10 Synthesis of DNA in one arm of the fork is in the same direction as fork movement.

Okazaki Fragments

Question: Figure 16-10 shows that the new chain synthesized in the lower arm 3′→5′ is in the direction of fork movement. How is this achieved?

DNA synthesis in the lower arm of the diagram (on the 5′→3′ template) is made in the direction *opposite* to fork movement and in short segments of 100–1,000 nucleotides in length. These are called *Okazaki* or *nascent fragments*. Thus replication of both arms can be represented as shown in Fig. 16-11.

The fragments are subsequently joined together. This overall pathway of assembly of new strands

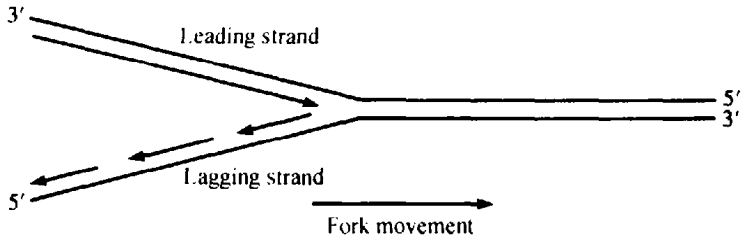


Fig. 16-11 Synthesis of DNA as nascent or Okazaki fragments in one arm of the replication fork.

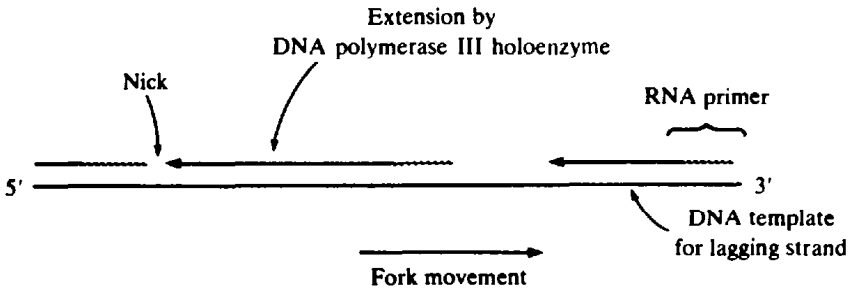


Fig. 16-12 Priming and growth of Okazaki fragments. The *nick* represents the unformed phosphodiester bond between the polynucleotide chains, in this case it is the RNA primer and the newly synthesized DNA segment.

at the replication fork is called *discontinuous DNA replication*. The strand that is made continuously in the direction of fork movement is called the *leading* strand and the other, the *lagging* strand.

Question: How are the Okazaki fragments initiated?

DNA polymerases are unable to initiate synthesis of a new chain, they can only *extend* a chain. A special type of RNA polymerase, called *primase*, makes short *RNA primers* opposite in direction to that of fork movement; the primers are made at many regions along the 5′→3′ template as it is progressively exposed by the unwinding of the helix. Following release of the primase after completing each primer, the DNA chain is extended backwards by DNA polymerase III holoenzyme. (The ability of the RNA polymerases to start chains de novo is a significant difference from the DNA polymerases.) Chain extension by DNA polymerase III continues until the newly synthesized DNA fragment comes up to the 5′ end of an RNA primer in the adjacent fragment, as shown in Fig. 16-12.

Question: What determines the sites of initiation of RNA primer formation along the 5′→3′ template?

RNA formation, from a DNA template, is initiated at sites on DNA known as *promoters* (Chap. 17). For primer formation at the moving replication fork the cell provides a mobile promoter that moves along the 5′→3′ template and functions at regular intervals to promote the initiation of RNA primer synthesis by the primase. The role of the mobile promoter (sometimes referred to as the *primosome*) on the *E. coli* chromosome is performed by the protein/enzyme DnaB. DnaB facilitates the action of primase (DnaG), with which it is presumably loosely associated, at regular intervals. (DnaB also has an important role in unwinding the DNA, see below.)

The RNA primer must be removed and replaced by DNA. This is accomplished as follows: The DNA polymerase III holoenzyme ceases its action (without dissociation from the DNA) when it reaches the 5′ end of the RNA primer, and DNA polymerase I takes over. A feature of DNA polymerase I, in contrast to other DNA polymerases, is its ability to effect replication at a nick. It doesn't matter if the polynucleotide *ahead* of the nick is DNA or RNA. Referring to Fig. 16-12, DNA polymerase I binds at the solid left arrowhead and continues to extend the DNA chain. At the same time, it removes the RNA primer through its 5′→3′ exonuclease activity. The overall effect is for it to replace the RNA by DNA, and this continues until the nick has shifted beyond the RNA section. This is an example of *nick translation*. Once the nick is bounded by DNA, it is sealed.

Question: How is a nick sealed in a DNA strand?

The enzyme used is *DNA ligase*. It catalyzes the formation of a phosphodiester link between adjacent 5′-phosphoryl and 3′-hydroxyl groups in double-stranded DNA. DNA ligase from *E. coli*

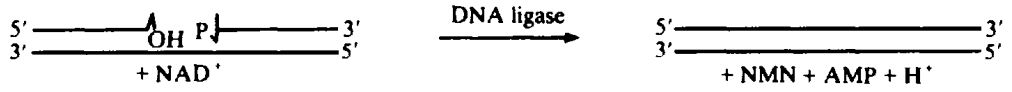


Fig. 16-13 The reaction catalyzed by DNA ligase in *E. coli*. (NMN = nicotinamide mononucleotide.) Eukaryotic ligases use ATP instead of NAD^+ as the coreactant.

requires NAD^+ as a coreactant, and the reaction mechanism involves the formation of an intermediate in which the adenyl group of NAD^+ is covalently attached to the enzyme. The 5′-phosphoryl terminus at the nick is activated by transfer of the adenyl group to form a *DNA-adenylate*. The phosphodiester bond is then formed by attack of the 3′ hydroxyl on the activated 5′ phosphate. The resultant reaction is shown in Fig. 16-13.

Unwinding the Double Helix

The two strands of DNA are wrapped around one another in the form of a helix, and the discussion so far has not addressed in detail the fact that these strands must *unwind* in order to expose the template for copying. The main route to unwinding DNA is through the action of *DNA helicases*. These are enzymes that couple the unwinding of double-helical DNA to hydrolysis of ATP. *E. coli* has several different helicases and one in particular, DnaB, has the major role to unwind the DNA during replication of the chromosome. DnaB moves along the 5′→3′ template strand at the leading edge of the replication complex. It is possible that unwinding is facilitated in other ways: by the binding of *single-stranded DNA binding protein (SSB)* to DNA ahead of the fork as well as by the introduction of *negative supercoiling* (Chap. 7) in this region by the enzyme *DNA gyrase*.

Question: What is the nature of the interaction of SSB with DNA?

SSB, also called *helix-destabilizing protein*, binds tightly and cooperatively to single-stranded DNA; i.e., when one molecule of SSB protein binds to the DNA, it *facilitates* the binding of more molecules of SSB. DNA is not a static structure; transient and localized strand separation occurs continuously. Thus, SSB, by binding cooperatively to the transient single-stranded regions, actually lowers the melting temperature of DNA; in other words, it aids DNA separation into component strands. As well as facilitating unwinding, SSB functions in another way. A single strand to which SSB protein is bound is a more rapidly processed template for DNA polymerase; this is an important factor in promoting efficient DNA chain growth.

Question: What is the mechanism of DNA gyrase action?

DNA gyrase belongs to a class of enzymes known as *topoisomerases*. Such enzymes, by catalyzing the concerted cutting and closing of strands in DNA, convert one topological form (*topoisomer*, Chap. 7) to another. There are two types of topoisomerases: *topoisomerase I* cuts just one strand of duplex DNA; *topoisomerase II* cuts across two strands of a DNA duplex. DNA gyrase is an example of a topoisomerase II.

Topoisomerase I catalyzes the *relaxation* of *negatively supercoiled* DNA; no coreactants are needed. Topoisomerase II catalyzes the negative supercoiling of DNA, using ATP as a coreactant, and it is also able to relax *positive supercoiling* (Chap. 7).

Topoisomeric forms of DNA are important in DNA replication in the sense that the introduction of negative supercoiling facilitates the unwinding of the fork. Also, in a closed molecule such as the bacterial chromosome, the unwinding of the fork causes positive supercoiling in the DNA ahead of it, and this hinders further unwinding. DNA gyrase has the ability to induce negative supercoiling as well as to relax positive supercoiling.

EXAMPLE 16.7

A model that summarizes the multitude of events involving the proteins and enzymes described in the foregoing discussion, is shown in Fig. 16-14. The leading strand is synthesized in a continuous manner by *DNA polymerase III holoenzyme*. Its action is facilitated by the binding of *SSB* to the single-stranded template made available through unwinding of the unreplicated duplex. This unwinding is accomplished by the *helicase* DnaB which is associated with the 5′→3′ template. The lagging strand is synthesized as short fragments (discontinuously). As DnaB moves along the 5′→3′ template it promotes the action of *primase* at regular intervals to make a short piece of RNA. This is extended by DNA polymerase III holoenzyme in the direction opposite to fork movement. When DNA polymerase III reaches the primer, which has initiated an adjoining fragment, *DNA polymerase I* takes its place and, through the process of nick translation, removes the RNA and replaces it by DNA. The nick that remains is sealed by *DNA ligase*. Finally, it is possible that the overall process is facilitated through the action of *DNA gyrase* in inducing negative supercoiling or relieving positive supercoiling in the DNA ahead of the fork.

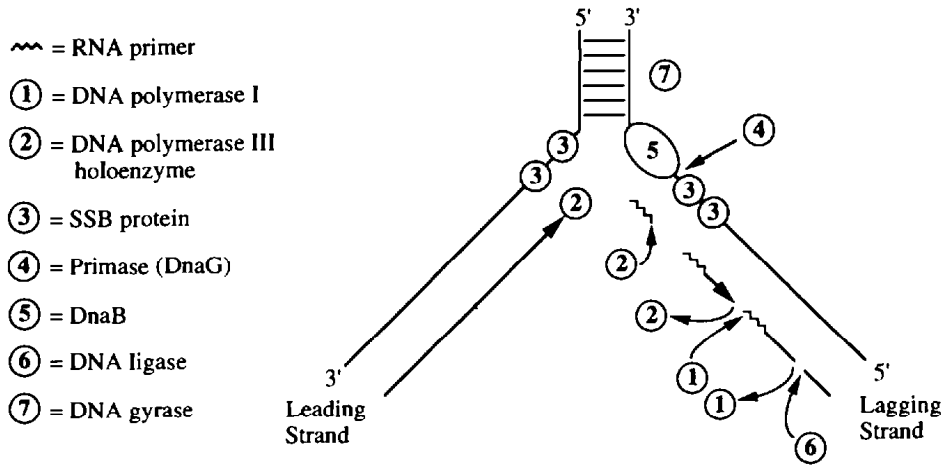


Fig. 16-14 Events involved in DNA chain growth at the replication fork.

EXAMPLE 16.8

DNA replication must be an accurate process. What happens if an incorrect base is incorporated by either of the DNA polymerases?

Mistakes in base incorporation can be made; this is largely a result of the transient existence of *tautomeric* forms of the bases (Chap. 7). If at the instant of insertion of a new nucleotide by DNA polymerase the base in the template shifts to its rare tautomeric form, which has altered base-pairing specificity, an incorrect nucleotide may be added to the chain; e.g., one containing guanine instead of adenine opposite the enol form of thymine.

The 3′→5′ exonuclease activity of DNA polymerase I, at least, functions to *proofread* for such mistakes. After the incorrect base is incorporated, it will not remain hydrogen-bonded to the tautomeric base in the template once the latter returns, almost immediately, to its more stable form. The 3′→5′ exonuclease activity shows a strong preference for a *frayed* or non-hydrogen-bonded end and removes the misincorporated nucleotide before chain growth proceeds further. DNA polymerase III holoenzyme also has the potential to proofread by the same mechanism.

16.6 MOLECULAR EVENTS IN THE INITIATION OF REPLICATION IN BACTERIA

The elongation phase of DNA replication in bacteria has been seen to involve many enzymes and proteins, and some are associated with discrete functional complexes, such as the DNA polymerase III holoenzyme. Initiation of replication also uses several proteins, and mutations in their genes have been very helpful in identifying these proteins.

EXAMPLE 16.9

Temperature-sensitive mutations, in particular, have been very valuable in helping to define many of the proteins involved in replication. Several of these proteins have already been discussed. Temperature-sensitive mutations take effect at a certain temperature, e.g., 42–47°C, and not at another, e.g., 30°C or less. Mutations that affect replication are called *dna* mutations. Many that have been identified in *E. coli* code for various proteins associated with DNA chain growth at the replication fork. For example, the gene *dnaG* codes for primase (the DnaG protein) which has already been discussed. Some, however, code for proteins involved also or exclusively with the *initiation of a cycle of replication* at *oriC*. Examples of these are *dnaA*, *B* and *C*.

Question: What is known about the *oriC* sequence in *E. coli* at which initiation is brought about by the action of the relevant proteins?

The *oriC* locus consists of a unique 245-bp sequence which contains a number of elements as shown in Fig. 16-15. These include four 9-bp segments called DnaA boxes which enable the *oriC* sequence to recognize and bind DnaA. Three 13-bp AT-rich segments at one end of *oriC* allow the DNA to "open up" during the process of initiation and provide an entry point for DnaB and DnaC. The *oriC* region and its immediate vicinity also contain numerous GATC sites which can be methylated. For simplicity they are not shown in Fig. 16-15. They have a role in preventing re-initiation of replication at *oriC* before the appropriate time in the cell cycle.

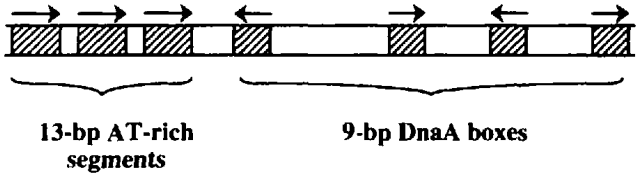


Fig. 16-15 Features of the 245-bp *oriC* sequence. The arrows show the orientation of the segment in each case.

Question: What proteins and other factors are involved in initiation at *oriC*?

Unlike the initiation of *Okazaki fragments* during elongation, initiation at *oriC* requires *RNA polymerase* (in contrast to primase it is sensitive to rifampicin, see Chap. 17), and DnaA, DnaB, DnaC, and the histone-like protein HU. The role of RNA polymerase is thought to be in bringing about *transcriptional activation* of *oriC*. This presumably facilitates the multiple molecular steps leading to successful initiation.

EXAMPLE 16.10

The cell embarks upon initiation at the appropriate stage of the cell cycle and this is probably triggered by the build-up of DnaA. DnaA binds ATP, and 10–20 monomers of DnaA-ATP, in conjunction with HU, bind to the DnaA boxes causing the *oriC* DNA to wrap around the DnaA-ATP multimer. This brings about melting (opening-up) of the segment spanning the three 13-bp AT-rich segments. DnaA then guides DnaB, from a complex with DnaC, into the melted region, and ATP is hydrolyzed to ADP. In the presence of SSB and DNA gyrase, the DnaB helicase unwinds the DNA bidirectionally. This provides the opportunity for entry of DNA polymerase III holoenzyme and DNA gyrase to commence DNA chain growth, thus generating the two (bidirectional) forks.

16.7 TERMINATION OF CHROMOSOME REPLICATION IN BACTERIA

The two replication forks generated at *oriC* must eventually come together and fuse to generate two complete new chromosomes. The forks meet in the *terminus region* (*tr*); there is a specific mechanism to ensure that they do not meet outside this region.

Question: What is the mechanism whereby fork meeting and fusion are restricted to the terminus region?

The terminus region of the *E. coli* chromosome covers about 25% of its length. Distributed over this region are six DNA terminator sequences (*TerA-F*), each of 23 bp. They are arranged as two opposed groups of three, as shown in Fig. 16-16. The terminators bind a *terminator protein* called Tus (for *terminus utilization substance*) with high affinity. The Tus-*Ter* complex can block fork movement when the fork approaches from one direction but not the other. Complexes with *TerB*, *C* and *F* are able to arrest the clockwise-moving fork; it is not arrested by those involving *TerA*, *D* and *E*. The latter arrest the anticlockwise-moving fork. Thus, the six terminators comprise a *replication fork trap* which allows a fork to enter the terminus region but not to leave it.

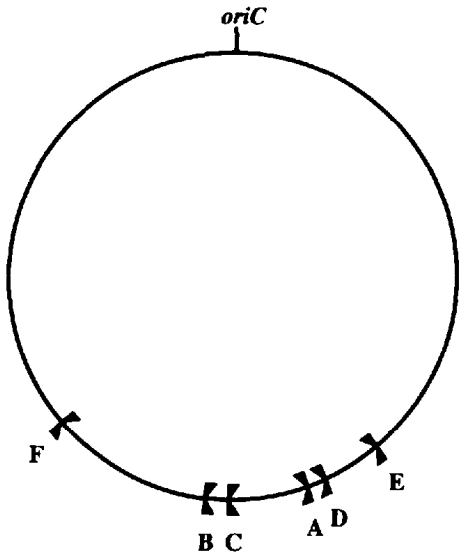


Fig. 16-16 Location of replication terminators on the *E. coli* chromosome.

Question: Why is it necessary to have six replication terminators?

On most occasions the forks meet and fuse in the vicinity of *TerC*. *TerC* appears to be the most frequently utilized terminator, probably because the clockwise fork would get to this site before the anticlockwise fork reached *TerA*. Presumably the terminators flanking the innermost ones function as back-ups when *TerA* or *C* fails.

Question: What is the advantage of restricting fork fusion to the terminus region?

The answer to this question is not clear. But it is known that upon completion of replication the two daughter chromosomes can remain linked together either as a covalent dimer or through intertwining in the terminus regions. Special enzymes (*recombinases* and a *topoisomerase*) are involved in the unlinking, which could be dependent also upon movement apart of the two chromosomes. It is possible that such unlinking is more efficient if the forks meet and fuse in the restricted terminus region.

16.8 INITIATION, ELONGATION, AND TERMINATION OF REPLICATION IN EUKARYOTES

Many of the proteins and enzymes involved in initiation at replication origins and DNA chain growth at replication forks have the same biochemical activities as their counterparts in bacteria. However, the situation regarding terminators and terminator proteins is less clear. Whether they exist to delineate to any extent individual replicons or clusters of replicons is not known. In the case of eukaryotic chromosomes, however, there is a special mechanism to replicate their ends which are known as *telomeres*.

Question: What are the structural features of origins of replication in eukaryotes and how is initiation achieved?

Most progress in understanding this has been made in the yeast *Saccharomyces cerevisiae* from the chromosomes of which sequences have been identified that allow initiation of replication when inserted into *plasmids* (see below). These sequences are called *autonomously replicating sequences* (ARSs). It is likely that some of them, at least, represent *bona fide* replication origins. Presumably the mechanism of initiation at these origins is similar to that which has been elucidated in bacteria.

Question: What is known about the process of DNA chain growth at replication forks in eukaryotes?

The structure of replication forks in eukaryotes is essentially the same as in bacteria. Chain growth is continuous on the leading strand and discontinuous on the lagging strand. There are equivalents of the polymerases, helicase, primase, SSB, etc., but there are clearly some differences. For example, two different polymerases, DNA polymerase δ and DNA polymerase α, function on the leading and lagging strand, respectively. Also, the mitochondrion has its own DNA polymerase.

Question: What is special about the ends of eukaryotic chromosomes and how are they replicated?

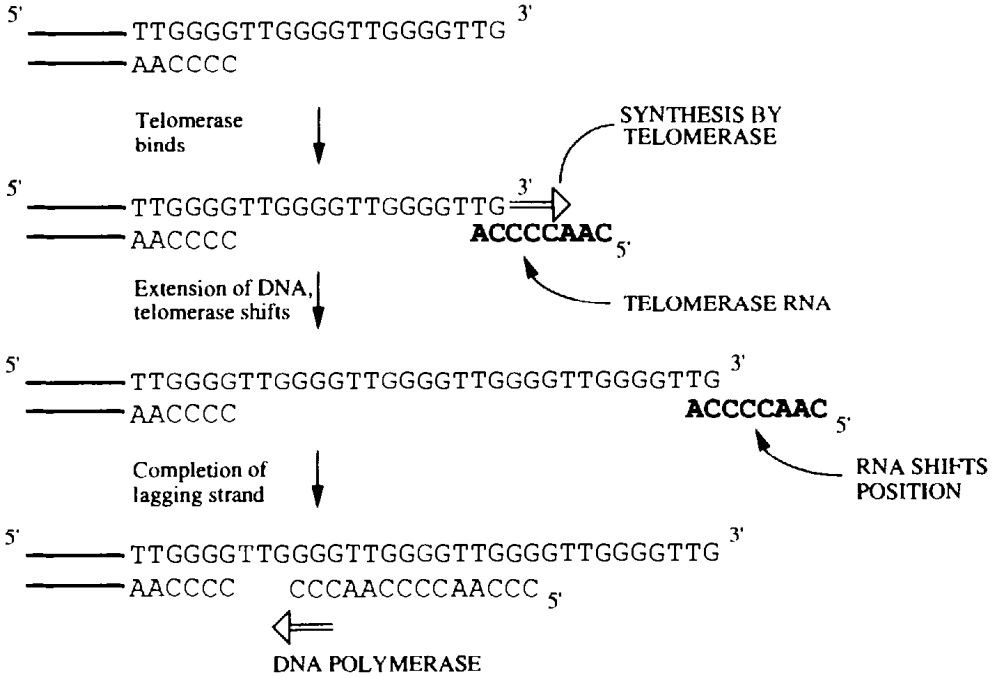


Fig. 16-17 Replication of the end of a eukaryotic chromosome by telomerase.

The ends of eukaryotic chromosomes (*telomeres*) have a special and unique structure. Because of the discontinuous manner of replicating the lagging strand at a replication fork, involving the formation and erasure of RNA primers, it is not possible for a fork to replicate right to the end of linear duplex DNA. The problem has been overcome through the existence of a repeated (hundreds) GC-rich hexameric sequence at the ends and a dedicated mechanism for its replication. A special enzyme, *telomerase*, contains an RNA molecule which provides a template for the elongation of the 5′→3′ strand, which is G-rich, as shown in Fig. 16-17. When the 3′ end is sufficiently elongated the lagging strand is presumably completed by DNA polymerase which contains primase as one of its subunits. Because of the large number of tandem repeats which make up the telomeres it is of little consequence if the precise number varies from one chromosome to another.

EXAMPLE 16.11

DNA in eukaryotes is tightly associated with histones in the nucleosome units (Chap. 7). Clearly, new nucleosomes must be assembled upon replication of the DNA. The histones are synthesized during S phase. The new DNA becomes packaged into nucleosomes immediately upon replication. There is some evidence to suggest that the parental histone octamers are transferred rapidly from in front of the advancing fork to the newly replicated DNA. Previously it was thought that the parental octamers were transferred exclusively to just one of the new daughter duplexes, but this is currently an unresolved issue.

16.9 INHIBITORS OF DNA REPLICATION

Inhibitors of DNA replication are sometimes valuable in the treatment of various types of disease. They fall into several classes:

Inhibitors of Nucleotide Biosynthesis

Examples of inhibitors of nucleotide biosynthesis are *methotrexate* and *fluorodeoxyuridylate*. They interfere with the production of dTTP, which has been discussed in Chap. 15.

Inhibitors that Interact with the DNA Template

Actinomycin D is a commonly used inhibitor of both DNA and RNA synthesis. Its planar structure binds noncovalently between the stacked base pairs of duplex DNA; this is called *intercalation*. In this situation the DNA functions as a poor template. Compounds that bind in a similar way include *acridine* and *ethidium*. These affect the fidelity of DNA replication.

Nucleotide Analog Inhibitors

A number of nucleotide analog compounds function by blocking further chain growth at the replication fork. The *2′3′-dideoxynucleosides* can be converted to the triphosphates (ddNTPs). In the case of bacteria, these are incorporated onto the 3′-hydroxyl end of a growing DNA chain, and because the new end now lacks a 3′ hydroxyl, no further additions can occur. They are used in conjunction with DNA polymerase I in the *dideoxy method* of Sanger for determining DNA sequences.

EXAMPLE 16.12

The dideoxy method of Sanger for sequencing DNA involves the copying of a single strand of the DNA by a DNA polymerase (DNA polymerase I was originally used) to yield new strands (radioactively labeled for identification purposes), which are terminated at certain positions through the incorporation of a dideoxy nucleotide at the 3′ end. A short primer, complementary to a sequence at one end of the single-strand template,

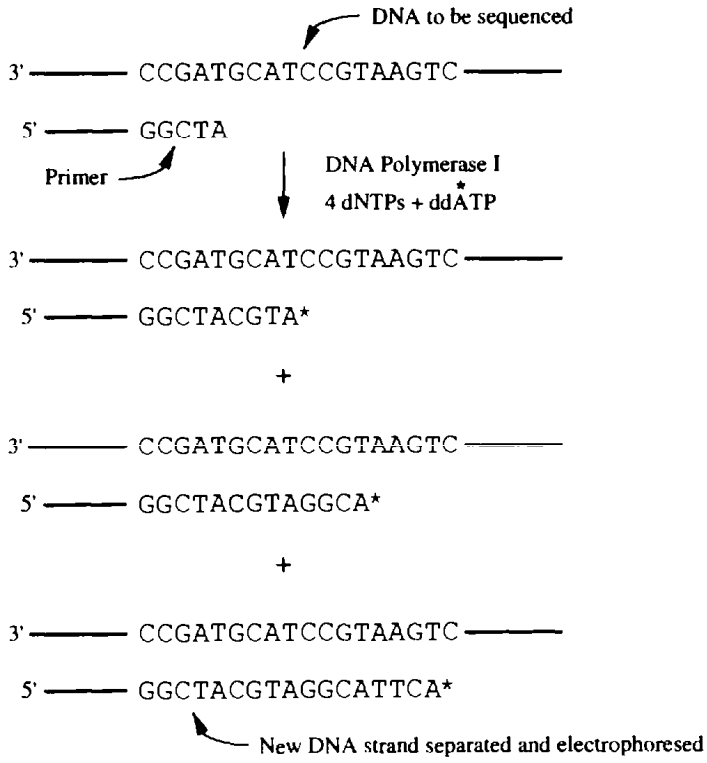


Fig. 16-18

is used to start synthesis from a fixed position. Termination at variable sites corresponding to the incorporation of a particular nucleotide is achieved by the inclusion of the ddNTP for that nucleotide. Four complete reactions (substrates plus enzyme) are set up, each containing, in addition, just one of the four ddNTPs. Thus, for a reaction using ddATP, the synthesis will proceed as in Fig. 16-18. The new strands of varying length are analyzed by gel electrophoresis (Chap. 4). From the collection of strand lengths produced by use of each of the four ddNTPs, the nucleotide sequence can be determined.

Inhibitors that Bind to Replication Proteins

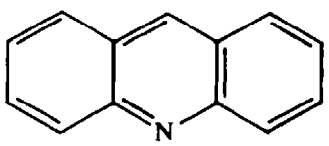
The *arylhydrazinopyrimidines* are potent inhibitors of DNA polymerase III from Gram-positive bacteria (*E. coli* is a Gram-negative organism). These compounds form *ternary complexes* with the polymerase and the DNA template.

Aphidicolin, a tetracyclic diterpenoid, is a potent inhibitor of mammalian nuclear DNA polymerases. It does not affect mitochondrial DNA polymerase.

Nalidixic acid and *novobiocin* bind to the A and B subunits, respectively, of *E. coli* DNA gyrase to inhibit its action and hence DNA replication.

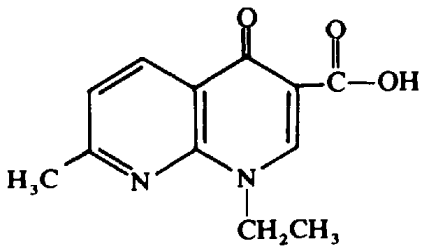
EXAMPLE 16.13

Acridine drugs have been used to treat malignant and parasitic diseases. The structure of acridine is:

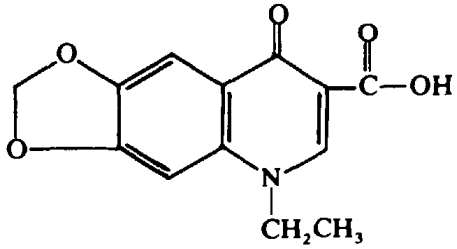


Acridine

Nalidixic acid is an example of a clinically useful *antibacterial* agent, but *oxolinic acid*, which acts in the same way, is 10 times more potent. The structures of these two antibiotics are:

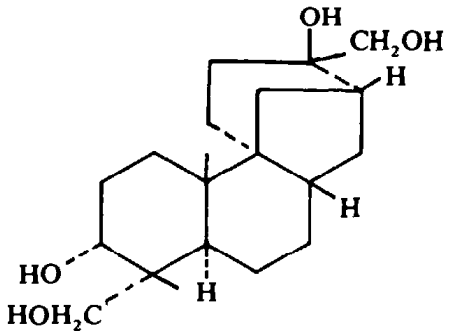


Nalidixic acid



Oxolinic acid

Aphidicolin has been used to treat herpes virus infections of the eye. Its structure is:



Aphidicolin

16.10 REPAIR OF DNA DAMAGE

Damage to DNA is caused by a number of agents, including uv irradiation, ionizing radiation, and various chemicals. Such damaged molecules can be deleterious or lethal to an organism, and a number of mechanisms exist for removing them. The best understood is that known as *excision repair* of damage caused by *uv irradiation.*

EXAMPLE 16.14

Upon uv irradiation pyrimidine bases in adjacent nucleotides of a DNA strand become covalently cross-linked. The cross-linking occurs mostly between two thymines to form a *thymine dimer* (Fig. 16-19). The thymine dimer causes a structural distortion within the DNA chain and represents a physical impediment to replication and transcription.

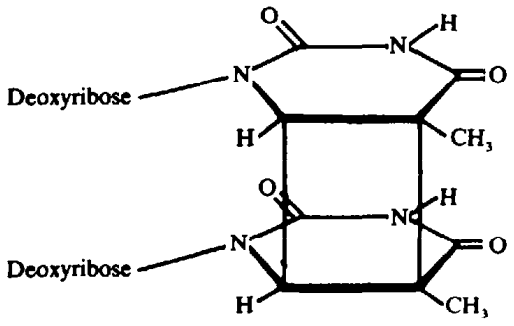


Fig. 16-19 Structure of a thymine dimer. The dimer is attached at two positions to the sugar, deoxyribose, and note its cyclobutane structure.

Question: How, if at all, is damaged DNA repaired?

In *excision repair* the damaged portion is cut-out and replaced by new DNA. The process in *E. coli* is well understood and summarized in Fig. 16–20. It occurs in four steps. First, a complex of three proteins, UvrABC, recognizes the damaged region and makes a single-strand cut on each side of the lesion, 12–13 nucleotides apart. The damaged segment is removed by the action of a helicase. The gap is filled by DNA polymerase I and the nick sealed by ligase.

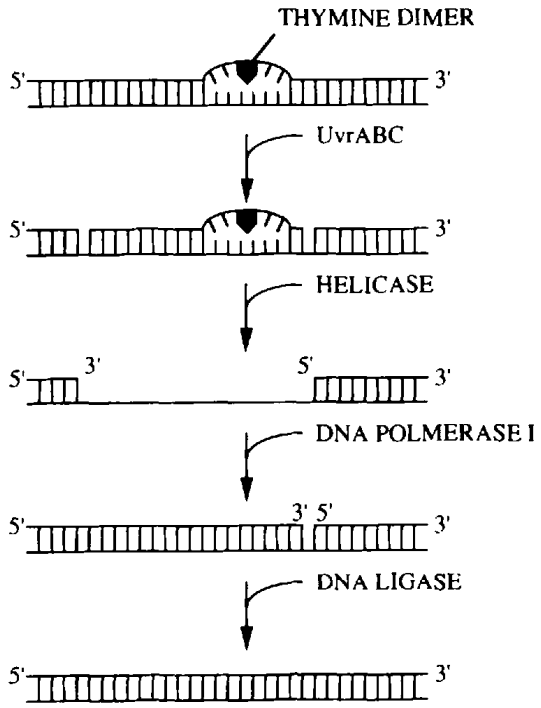


Fig. 16-20 Events in excision repair of uv-induced damage to DNA.

EXAMPLE 16.15

Excision repair in mammalian cells is considered to occur by a similar mechanism. Inherited defects in this repair pathway can lead to the disease *xeroderma pigmentosum*, in which the skin is very sensitive to sunlight and results in a high incidence of skin cancer.

Question: Is excision repair the only means of correcting uv-induced damage to DNA?

No. Dimers can also be repaired directly by enzymatic *photoreactivation*. The photoreactivating enzyme binds to the DNA in the region of the dimer to form a complex that absorbs visible light and catalyzes cleavage of the covalent linkage between the components of the dimer. Photoreactivation occurs in both bacterial and mammalian cells.

16.11 RECOMBINANT DNA AND ISOLATION OF GENES

With the development of *recombinant DNA technology*, it has been possible to isolate particular segments of DNA, sometimes containing all or part of a gene, and examine their structure and other

properties in detail. This has led to very significant advances in understanding, at a very refined level of molecular detail, many aspects of replication and expression of the genetic material.

Recombinant DNA is an artificially constructed molecule containing DNA segments from different organisms. If one of the segments carries an *origin of replication*, the recombinant, or *joint*, molecule has the potential to be replicated in an organism that is able to recognize that origin. In such a situation, the segment carrying the origin, which enables replication, is referred to as a *vector*.

EXAMPLE 16.16

There are several methods available for making recombinant DNA molecules. One of the most commonly used approaches takes advantage of the single-stranded ends of DNA segments generated by *type II restriction endonucleases* such as *Eco*RI. Fragments produced by this enzyme contain overlapping 5′ termini (Chap. 7). These short single-stranded termini spontaneously join to one another under suitable conditions of temperature and ionic strength, through complementary base pairing. This process is called *annealing* (Fig. 16-21). The rejoined molecule is held together through four base pairs, but there is a nick on each strand. These can be sealed by DNA ligase.

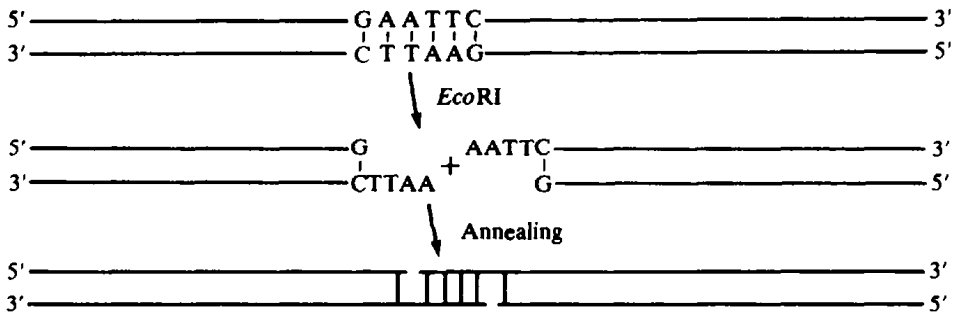


Fig. 16-21 Annealing of DNA through base pairing of overlapping termini generated by *Eco*RI.

Question: How is a segment of DNA linked to a suitable vector?

Commonly used vectors are *bacterial plasmids*. These are small circular DNA molecules of 3–100 kb. Circularity is essential for their replication. If a plasmid is enzymatically *cut* at only one site by, say, *Eco*RI, the linear fragment obtained can be annealed with a fragment that has been formed by cutting other DNA with *Eco*RI. This generates a circular joint molecule, which is sealed by DNA ligase (Fig. 16-22). The circular joint molecule is introduced into an appropriate host, such as *E. coli*, through a process known as *transformation*; this involves treatment of the bacterial cells with $CaCl_2$ at low temperature (to allow uptake of DNA), followed by incubation with the DNA under appropriate conditions. The foreign DNA replicates under the control of the vector. If the foreign DNA carries an intact gene, it is possible that this gene will be expressed in the bacterium. This type of approach has led to the mass production of human insulin in bacterial cells. The overall technique is sometimes referred to as *gene cloning*.

16.12 THE POLYMERASE CHAIN REACTION

Question: How is it possible to obtain sufficient quantities of a *particular* DNA sample to enable a determination of its properties?

This can be achieved using the polymerase chain reaction (*PCR*) which is indeed a type of chain

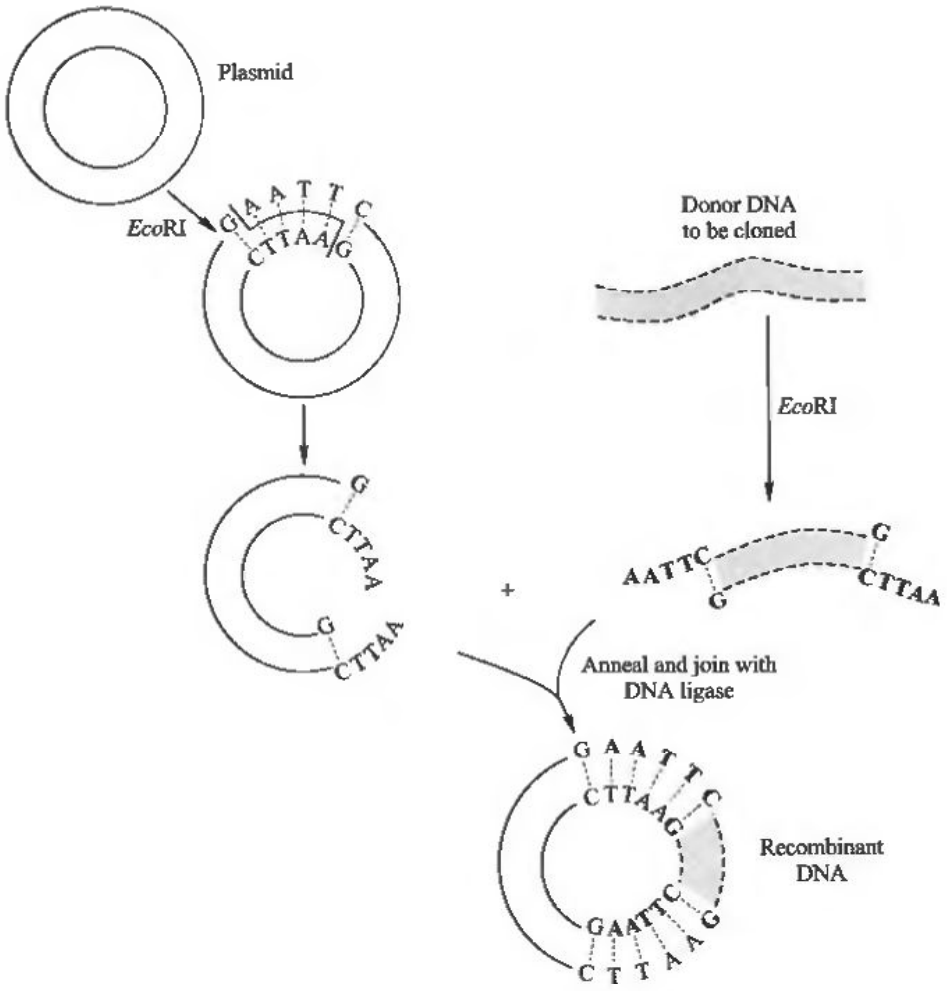


Fig. 16-22 Steps in the formation of a joint, or recombinant, DNA molecule.

reaction. PCR provides a way to generate a vast number of double-stranded copies of a specific DNA sequence. These copies are typically ~1 kb long, but by specially modifying the reaction longer products can be made.

The polymerase chain reaction uses: (1) a thermostable DNA polymerase, such as *Taq polymerase* derived from the bacterial thermophile *Thermus aquaticus*, (2) a *DNA template* which is to be amplified, (3) *two primers*, each typically of around 20 nucleotides, which anneal to distinct parts on the complementary strands of the target and serve as sites for commencing DNA polymerase action, (4) a solution including the four deoxynucleoside triphosphates dATP, dCTP, dGTP and dTTP, Mg^{2+}, salts and pH buffer.

PCR uses DNA polymerase to make complementary copies of DNA corresponding to the region of interest. To direct the enzyme to the correct sequence, PCR relies on primers which anneal to complementary sequences in target single-stranded DNA. DNA polymerase cannot start DNA synthesis *de novo*, and can extend the chain only from the annealed primers. An excess of these primers is provided in the reaction mixture to allow sufficient material for generating amplified product.

PCR reactions include the cyclical use of high temperatures which can lead to the denaturation of thermolabile enzymes; for this reason a thermostable DNA polymerase, Taq polymerase, is typically used.

For PCR, there are only three steps involving DNA. These are (1) double-stranded DNA denaturation, (2) DNA annealing and (3) DNA extension. Each step is usually conducted at a

Triacylglycerol has no polar interaction with the membrane phospholipids and is either released into the cytosol as tiny lipid droplets or into the lumen of the ER. In fat cells, oil droplets in the cytosol coalesce, migrate toward and fuse with the large central oil droplets. In the liver and intestine, triacylglycerol is packaged into lipoproteins (VLDL and chylomicrons, respectively), which then are secreted into the circulation (see Fig. 13-11).

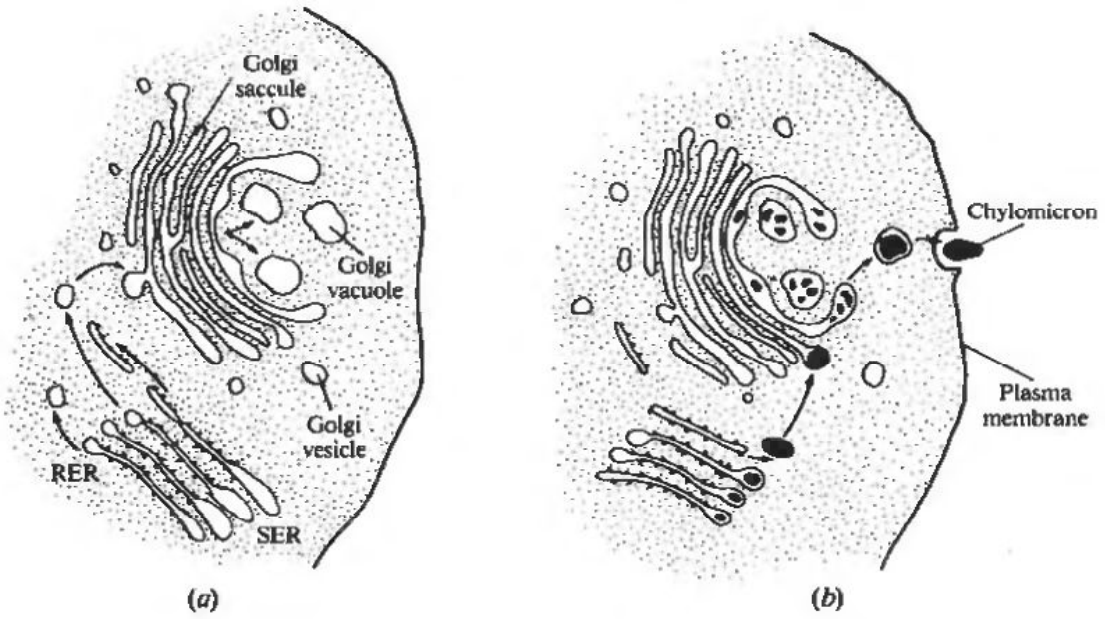


Fig. 13-11 Secretion of chylomicrons and VLDL: (*a*) from rats fasted for 24 h; (*b*) from rats 15–60 min after fat feeding.

13.8 SYNTHESIS OF PHOSPHOLIPIDS AND SPHINGOLIPIDS

Phospholipids

Phosphatidylcholine, a major phospholipid constituent of membranes and lipoproteins, is synthesized de novo in liver cells. The synthesis occurs on the ER and is linked, through 1,2-diacylglycerol, with the synthesis of triacylglycerol. Three compounds specifically involved in the synthesis of phosphatidylcholine are:

1. Choline

$$HO\text{-}CH_2\text{-}CH_2\text{-}N^+(CH_3)_3$$

2. Choline phosphate

$$^{2-}O_3P\text{-}O\text{-}CH_2\text{-}CH_2\text{-}N^+(CH_3)_3$$

3. Cytidine diphosphocholine (CDP-choline)

$$\text{Cytidine-O-}\overset{\overset{O}{\|}}{\underset{\underset{O^-}{|}}{P}}\text{-O-}\overset{\overset{O}{\|}}{\underset{\underset{O^-}{|}}{P}}\text{-O-}CH_2\text{-}CH_2\text{-}N^+(CH_3)_3$$

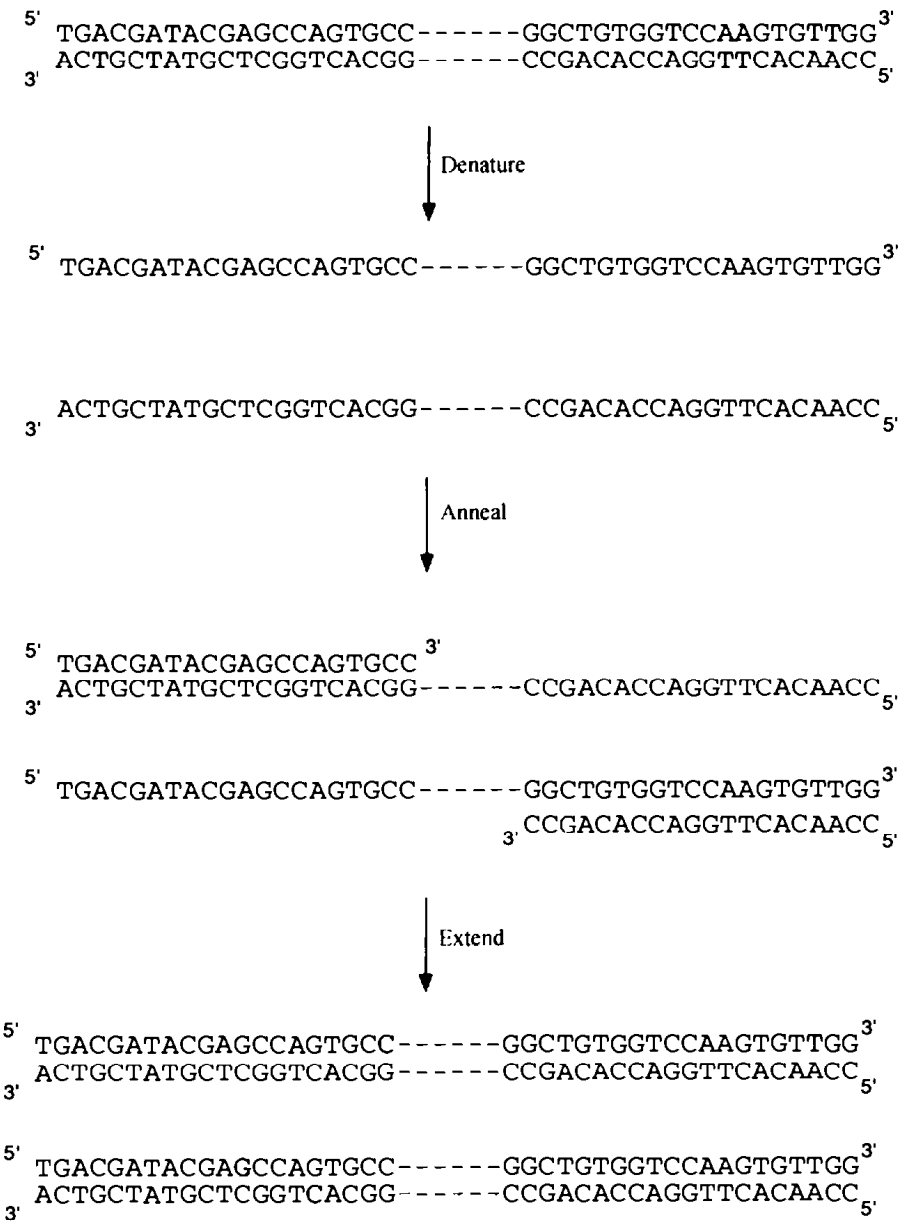


Fig. 16-23 Denaturation, annealing and extension processes in a PCR.

Question: Can amplification by PCR occur if the primers cannot bind to a target DNA sequence?

The annealing of primers is extremely important because this stage determines the specificity of the PCR reaction and defines the DNA sequence to be amplified. If neither primer can bind, DNA polymerase cannot extend DNA and, as a consequence, no DNA is synthesized. If only one primer binds, amplification is not exponential. For successful PCR to occur, both primers must anneal.

Question: Why is a thermostable DNA polymerase used in PCR?

Cycling through a range of high temperatures would denature many common laboratory enzymes. This is particularly the case when the highest temperatures (~95°C) are used to separate the DNA strands. If a DNA polymerase were not thermostable it would denature. Thermostable DNA polymerases can be derived from organisms such as the bacterium *Thermus aquaticus* which grows at high temperatures.

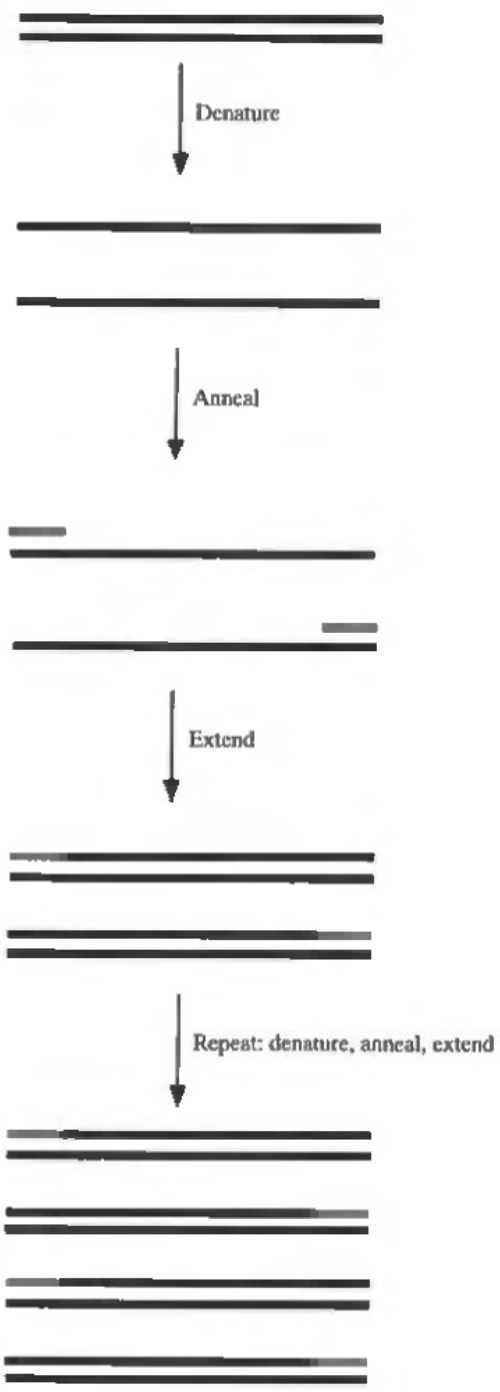


Fig. 16-24 The PCR reaction.

Solved Problems

SEMICONSERVATIVE REPLICATION OF DNA

16.1. Define the term *template* as it applies to DNA replication.

SOLUTION

The template is the DNA strand that binds directly to the various replication enzymes and that defines

the sequence of the newly synthesized DNA strand. The DNA duplex cannot be copied per se. Rather, it must unwind into its component single strands, in which the nucleotide sequence is accessible to copying through base pairing.

16.2. Consider a linear duplex DNA molecule that undergoes replication through five successive generations. What would be the proportion of the *original* DNA within the *total* DNA population?

SOLUTION

The amount of DNA doubles in each generation. Thus, five successive generations result in a 32-fold (2^5) increase in the DNA. Therefore, the original DNA represents 3.125 percent of the total.

16.3. With respect to the situation outlined in Prob. 16.2, how is the original DNA molecule distributed among the progeny molecules?

SOLUTION

The progeny consists of 32 molecules. The two single strands of the original (or parental) DNA remain intact. Each makes up half of two of the progeny molecules.

TOPOLOGY OF DNA REPLICATION

16.4. A replicating bacterial chromosome exists as a so-called *θ structure* (because of its appearance in electron micrographs and diagrammatically), in which the upper closed section represents the duplicated portion, or replication bubble. Why is there now an origin of replication situated at the middle of *each* arm of the bubble?

SOLUTION

Replication of the bacterial chromosome commences at a unique site, the origin, and replication forks diverge from this site at approximately equal rates. Consequently, the origin is duplicated completely and each daughter origin remains at the central portion of the expanding bubble. This is in contrast to unidirectional replication, which would leave the origin only partially replicated and located at the stationary fork as the replication fork moved away.

16.5. What is the major difference, with respect to the topology of replication, between circular bacterial and linear eukaryotic chromosomes?

SOLUTION

The major difference arises from the existence of only *one* origin of replication in the bacterial chromosome compared with a *multitude* of origins in a eukaryotic chromosome. This gives rise to a single replication bubble in the former, and clusters of tandemly arranged bubbles in a eukaryotic chromosome. In the case of bacteria, the two replication forks that meet at termination arise from the same origin; in eukaryotes they arise from neighboring origins.

CONTROL OF DNA REPLICATION

16.6. What conditions give rise to *multiforked chromosomes* in bacteria?

SOLUTION

A multiforked bacterial chromosome is one that contains more than two replication forks and results from *reinitiation* at the daughter origins within a replication bubble. In this situation, cycles of replication are completed at more frequent intervals, and this gives *shorter* generation times. It occurs under conditions of fast growth, induced by nutritionally rich media.

16.7. Under what conditions do sections of eukaryotic chromosomes exist as *multiforked* structures?

SOLUTION

Unlike the situation in bacteria, multiforked structures in eukaryotic chromosomes do *not* form in response to a demand for increased rate of DNA replication. But, multiforked DNA does appear over restricted regions of eukaryotic chromosomes; this gives DNA (or gene) amplification and occurs in specific situations.

16.8. Chromosome replication in bacteria can be divided into three stages. Which of these stages is central to the overall control of the process?

SOLUTION

The stages of DNA replication in bacteria are *initiation*, *elongation*, and *termination*. Control of replication is effected at the level of initiation. Thus, the frequency of initiation determines the frequency of completion of cycles of replication. The rate of replication-fork movement (elongation) remains fairly constant under conditions that change the overall rate of replication. There is no evidence to suggest that termination has any control over the rate of replication.

ENZYMOLOGY OF DNA REPLICATION

16.9. Distinguish between the *substrates* and *template* of DNA polymerase.

SOLUTION

The substrates undergo chemical modification and include the four deoxynucleoside triphosphates. The template undergoes no alteration but provides the instructions for the order of assembly of nucleotides into the growing chain. The template for DNA synthesis is the parental, single-stranded DNA.

16.10. Do DNA polymerase I and III of *E. coli* extend a DNA chain by different mechanisms?

SOLUTION

No. All DNA polymerases use the deoxynucleoside 5′-triphosphates to add new nucleotide units, one at a time, onto the 3′-hydroxyl terminus of the growing chain. The main differences between DNA polymerase I and III relates to the type and relative contribution each makes to replication and repair-synthesis of DNA.

16.11. DNA polymerase I of *E. coli* is a multifunctional enzyme. What is meant by this?

SOLUTION

DNA polymerase I, which consists of a single polypeptide chain, contains three distinct active sites. These are polymerase, $5' \rightarrow 3'$ exonuclease, and $3' \rightarrow 5'$ exonuclease. Each activity has an important biological role. The $5' \rightarrow 3'$ exonuclease activity can be separated from the other two by proteolytic cleavage of the enzyme.

16.12. What constitutes the DNA polymerase III core enzyme of *E. coli*?

SOLUTION

DNA polymerase III occurs within the cell as a functional complex of 10 polypeptide chains. This is called the *holoenzyme*. A subcomplex containing three of these chains (α, ε, and θ) is readily isolated and exhibits polymerase activity. It is called the DNA polymerase III core enzyme.

16.13. What is understood by *discontinuous* DNA replication?

SOLUTION

DNA chain growth occurs on both daughter arms at a replication fork. On one arm, chain growth occurs continuously ($5' \rightarrow 3'$), in the same direction as fork movement. On the other arm, chain growth occurs in separate short pieces ($5' \rightarrow 3'$) and in the direction opposite to fork movement. The short pieces (nascent or Okazaki fragments) subsequently join. Replication in the latter fashion is known as *discontinuous* DNA replication.

16.14. What is meant by the terms *leading* and *lagging strands* in DNA replication?

SOLUTION

The leading strand is that which is synthesized continuously and in the same direction as fork movement. The lagging strand is synthesized discontinuously, i.e., in short pieces that are subsequently joined, and in the direction opposite to fork movement.

16.15. What is the role of the $3' \rightarrow 5'$ exonuclease of DNA polymerase I of *E. coli*?

SOLUTION

This enzyme has a proofreading role. At a low random frequency, incorrect bases (in the form of nucleotides) are inserted into the growing DNA chain. This results from the existence of rare tautomeric forms of the four bases, which, if occurring transiently in the template at the moment of insertion of an incoming nucleotide, will cause a mistake in base pairing. When such a template nucleotide shifts back to its preponderant form, a base pair mismatch results. The $3' \rightarrow 5'$ exonuclease recognizes the mismatch and catalyzes the hydrolytic removal of the nucleotide from the end of the chain before elongation resumes.

16.16. A characteristic feature of the DNA polymerases is their inability to initiate synthesis of a polydeoxynucleotide chain; they can only *extend* an existing chain. How is initiation of the nascent fragments within the discontinuously synthesized DNA strand achieved?

SOLUTION

Unlike the DNA polymerases, RNA polymerase is able to initiate a new RNA chain, using DNA as a template (Chap. 17). The DNA polymerases are able to extend the DNA from an RNA primer. In discontinuous DNA chain growth, a particular type of RNA polymerase, called *primase* in *E. coli*, lays down short RNA primers at fairly regular base intervals, as unwinding of the helix at the replication fork proceeds. These primers are involved in the initiation of synthesis of nascent DNA chains by DNA polymerase.

16.17. In *E. coli*, what is the function of the $5' \rightarrow 3'$ exonuclease activity of DNA polymerase I in the overall replication process?

SOLUTION

DNA polymerase I has an important role in the assembly of new DNA chains via discontinuous DNA synthesis in the lagging strand at a replication fork. DNA polymerase III is responsible for the bulk of the synthesis of each nascent DNA fragment. When chain growth by this enzyme reaches the 5′ RNA-primer end of an adjoining fragment of DNA, polymerase I takes over and continues DNA extension through the RNA-primer region. In this case the segment of RNA is removed through the $5' \rightarrow 3'$ exonuclease activity of DNA polymerase I.

16.18. What reaction is catalyzed by DNA ligase, and what is its role in DNA replication?

SOLUTION

DNA ligase catalyzes the covalent linkage of two segments of DNA. A phosphodiester link is formed between adjacent 5′-phosphoryl and 3′-hydroxyl groups within *duplex* DNA. In other words, DNA ligase

is able to *seal* a nick. DNA ligase has an important role in several reactions involving DNA. In replication it functions to join the nascent DNA fragments of the lagging strand. This follows replacement of the RNA primer by DNA through the action of DNA polymerase I.

16.19. What are topoisomerases, and how might they be involved in the process of DNA replication?

SOLUTION

Topoisomerases catalyze the conversion of one topological form of DNA to another. There are two general enzyme types, I and II. The former acts by cutting just one strand of duplex DNA, while type II cuts across both strands of DNA. DNA gyrase, from bacteria, is an example of a type II topoisomerase and is known to be involved in DNA replication. It can induce negative supercoiling in DNA as well as being able to relax positive supercoiling. It is required for initiation of a cycle of replication in bacteria. In this case, it probably introduces negative supercoiling at the origin, thus allowing the binding of enzymes and other factors involved in initiation. Also, it facilitates the unwinding of the duplex at the fork by inducing negative supercoiling ahead of the fork or by relaxing the positive supercoiling that may build up in this region.

INITIATION OF REPLICATION IN BACTERIA

16.20. What is understood by the term *initiation of a cycle of replication* in bacteria?

SOLUTION

DNA replication in bacteria is a cyclic process in the sense that at regular time intervals, depending on the richness of the growth medium, initiation of replication is effected at the chromosome origin (*oriC*). A number of specific proteins and enzymes are directly involved in the initiation of replication, and they are not involved in the subsequent growth of the DNA. Once initiated, a cycle of replication runs to completion in the absence of initiation proteins or enzymes; but in the absence of these enzymes and factors, no new cycles start.

16.21. Distinguish between the roles of *RNA polymerase* and *primase* in the DNA replication process in bacteria.

SOLUTION

Both of these enzymes synthesize RNA. RNA polymerase is inhibited by the drug rifampicin; primase is not. Initiation of a cycle of replication, from *oriC*, is blocked by rifampicin, thus showing that RNA polymerase is required for initiation. Primase, on the other hand, is involved in initiation of nascent fragments in the lagging strand during the subsequent elongation or chain-growth stage. This stage is unaffected by rifampicin. Thus, these two enzymes, both of which catalyze the synthesis of RNA, have distinct roles in the process of DNA replication.

TERMINATION OF REPLICATION IN BACTERIA

16.22. What is meant by the term *replication fork trap*?

SOLUTION

Toward the end of a round of replication of the circular bacterial chromosome the two replication forks approach one another in the terminus region. It is important that they do not meet outside this region. Spread over this region are six polar terminators (DNA sequences) which bind a terminator protein and arrest a replication fork. The terminators are organized as two opposed groups of three such that a replication fork can enter the region but not leave. This organization of terminators over the terminus region provides a replication fork trap.

INITIATION, ELONGATION, AND TERMINATION OF REPLICATION IN EUKARYOTES

16.23. Why is there a greater variety of *DNA polymerases* in mammalian cells than in bacterial cells?

SOLUTION

As in bacterial cells two different DNA polymerases (δ and α) are involved in nuclear DNA replication. However, a distinct DNA polymerase (β) is involved in the repair of nuclear DNA, while mitochondria have their own polymerase for replication of their own DNA.

16.24. Why do eukaryotes contain the enzyme *telomerase*, while it is absent from *E. coli*?

SOLUTION

Telomerase is an enzyme that has a specific role in replicating and maintaining the integrity of the ends of linear chromosomes in eukaryotes. Such ends are absent in DNA from *E. coli*.

INHIBITORS OF DNA REPLICATION

16.25. Why does 2′,3′-ddCTP inhibit DNA polymerase I?

SOLUTION

DNA polymerase I catalyzes the addition of the monophosphate from 2′,3′-ddCTP onto the 3′-hydroxyl end of the growing chain and opposite guanosine on a template DNA strand. The growing chain is then devoid of a 3′ hydroxyl at the terminus, and the addition of further nucleotide units is prevented.

16.26. Would 2′,3′-ddCTP block the replication of $d(AT)_n$ by DNA polymerase I?

SOLUTION

$d(AT)_n$ is a polydeoxynucleotide containing alternating A and T residues in each strand of a duplex structure. 2′,3′-ddCTP has no effect on its replication by DNA polymerase I because G residues, to which this chain-terminating inhibitor would base-pair, are absent from the template.

REPAIR OF DNA DAMAGE

16.27. Distinguish between *replication* and *repair synthesis* of DNA.

SOLUTION

Replication refers to the reproduction, through the copying of template DNA strands, of complete chromosomes. In this case DNA synthesis is extensive. *Repair synthesis* of DNA is involved in correcting the damage (caused by physical or chemical treatments) within isolated portions of DNA. In this case, DNA synthesis is restricted to the immediate vicinity of the damage and is usually only minor in extent.

16.28. What enzymes are involved in removing the damaged portion of DNA during excision repair of uv-induced damage in bacteria?

SOLUTION

Four enzymes are directly involved. They are the UvrABC endonuclease, helicase, DNA polymerase I, and DNA ligase. The first cleaves the DNA strand on both sides of the pyrimidine dimer. The second enzyme removes the single-strand nucleotide segment containing the pyrimidine dimer. DNA polymerase I fills the gap with new DNA, and the nick is sealed by DNA ligase.

RECOMBINANT DNA AND ISOLATION OF GENES

16.29. What features of a *recombinant* plasmid DNA molecule are essential for its replication in a host cell?

SOLUTION

First, the recombinant plasmid DNA must carry an origin of replication that is *functional* in the particular host. Second, the molecule must be *circular*, or be capable of circularization.

16.30 A piece of human DNA carrying a particular protein-encoding gene has been *cloned* in a bacterial cell. Will the human gene be *expressed*?

SOLUTION

The human gene will be replicated, along with the *vector* in the recombinant DNA, but it won't necessarily be expressed. This only occurs if its disposition within the recombinant molecule is such that it is transcribed as mRNA and then translated into protein. For *transcription*, the gene must be positioned appropriately with respect to a bacterial *promoter*. For successful translation, the gene must be free of introns (Chap. 17), and its transcript must contain a bacterial ribosomal binding site (Chap. 17) at the correct location.

THE POLYMERASE CHAIN REACTION

16.31. Why is it best if the two primers in a PCR reaction mixture have similar G+C contents?

SOLUTION

The specificity of a PCR depends upon precise recognition of complementary regions on target-DNA strands. One way this specificity can be adjusted is by increasing the temperature of annealing, so only regions of high complementarity can bind. If the primers differ substantially in their G+C content, they will have different melting temperatures, and thus different stringency requirements.

16.32. How can PCR be used to introduce mutations into a DNA sequence?

SOLUTION

Errors can be introduced during the extension phase of a PCR because the thermostable DNA polymerase has a specific fidelity governed by the quality of the enzyme and any "proofreading" activity. Apart from this situation, mutations can be deliberately engineered into the product by incorporating mismatched bases into one or both primers. During the first cycle of the PCR these will be slightly mismatched, but subsequent products will incorporate the changed sequence as a simple consequence of semi-conservative DNA synthesis.

16.33. What are common mistakes people make when they are learning about PCR?

SOLUTION

People often confuse the primers. Two primers are needed and these are typically different sequences so they can bind to opposite strands. Also, these primers are present in large amounts in solution (in excess) to provide sufficient primers for each reaction. Each PCR product has, at its 5′ end, a primer which was originally supplied.

Supplementary Problems

16.34. (*a*) What is meant by *unidirectional* replication of DNA? (*b*) Is the rolling-circle mechanism of replication unidirectional or bidirectional?

16.35. The DNA synthesis (S) phase of the cell cycle is very short (< 20 min) in embryonic cells of some organisms. What process takes place that allows such rapid replication of all the chromosomal DNA? (*Hint*: Consider the size of the replicons.)

16.36. All three enzymatic activities of DNA polymerase I of *E. coli* are utilized in the overall replication process. Describe these.

16.37. Priming of DNA chain growth occurs at more than one fixed position on the bacterial chromosome. What is that position?

16.38. By what means can DNA gyrase assist in unwinding DNA at the replication fork?

16.39. Is DNA helicase a topoisomerase?

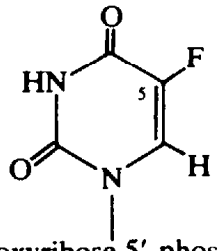
16.40. What structural features of the ends of DNA chains at a nick are necessary for sealing by DNA ligase?

16.41. What is meant by the proofreading function of DNA polymerase?

16.42. Devise a possible sequence of protein- (enzyme)-mediated events for the generation of a replication bubble at an origin of replication in double-helical DNA.

16.43. What is meant by *dna mutation*?

16.44. Would fluorodeoxyuridylate be expected to block initiation of DNA chain growth? Its structure is



16.45. What effect would an inhibitor of RNA polymerase be expected to have on DNA replication in bacteria?

16.46. Irradiation by uv light can cause the formation of thymine dimers. Does this bring about cross-linking of complementary strands of double-helical DNA?

16.47. What enzymes are common to the processes of replication of DNA and excision repair of DNA damage in *E. coli*?

16.48. Is excision repair the only means of removing thymine dimers from DNA?

16.49. What feature of circular, double-helical DNA is essential for it to function as a vector for DNA cloning?

16.50. What plays a more important role in determining the selection of which region is amplified during PCR: the primers or the DNA polymerase? Why?

Chapter 17

Gene Expression and Protein Synthesis

17.1 INTRODUCTION

Most genes are ultimately expressed as protein (Chap. 16). The process by which this is accomplished is called gene expression. In this process a sequence of deoxynucleotides in DNA (which defines the gene) is first *transcribed* into a sequence of ribonucleotides in RNA (*messenger RNA*, or *mRNA*). This is then *translated* into a sequence of amino acids to give a polypeptide of defined length. The amino acid sequence within the latter determines the manner in which the molecule folds upon itself to yield the biologically active protein.

In bacterial cells, there is no membrane surrounding the DNA *nucleoid*, and both *DNA transcription* and *RNA translation* proceed within the single cell compartment. In eukaryotes, the nucleus is bounded by a membrane (Chap. 1). Transcription occurs within the nucleus, and the mRNA must pass into the cytoplasm, where it is translated. Frequently, the immediate polypeptide product of translation is subsequently modified, sometimes in a process that enables it to be transported out of the cell in which it is made.

The sequence of nucleotides within the *single-stranded* mRNA is assembled according to the complementary-base-pairing (Chap. 7) instructions from one of the strands of duplex DNA, which contains the gene. The DNA strand that bears the same sequence as the mRNA (except for T instead of U) is called the *coding strand* or *sense strand*. The other strand of DNA which acts as the template for transcription is called the *template* or *antisense* strand. Some textbooks do not define "sense" and "antisense" in the way described here, and for this reason it may be preferable to use "coding" and "template" when referring to a particular strand.

The *genetic code* is the basis for converting a nucleotide sequence of mRNA into an amino acid sequence of a polypeptide. It is the genetic code that describes how various *combinations* of nucleotides (of which there are only four types in DNA or RNA) can be read as individual amino acids (of which there are 20 types). The nature of the genetic code was elucidated in the 1960s.

17.2 THE GENETIC CODE

Because there are 20 amino acids and only four nucleotides, there must be a combination of at least three nucleotides to define each amino acid. A code based on two nucleotides would provide only 4^2 or 16 combinations, which is insufficient. Proof that the *codon* for each amino acid consists of three nucleotides was provided by genetic studies of the effects on the polypeptide product of nucleotide addition to or deletion from a gene.

Question: A trinucleotide-based code would provide 4^3 or 64 codons. Are the extra codons used?

Yes, it turns out that they are all used. In the vast majority of cases, a single amino acid has more than one codon. For this reason the code is said to be *degenerate*.

EXAMPLE 17.1

Degeneracy of the code is very obvious from an examination of the codon assignments shown in Table 17.1. For example, there are six codons for leucine. It should be noted that the nucleotide components of each triplet

Table 17.1. Codon–Amino Acid Assignments of the Genetic Code

First Position	Second Position				Third Position
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	(CT)*	(CT)*	A
	Leu	Ser	(CT)*	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met(CI)†	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val(CI)†	Ala	Glu	Gly	G

*CT = chain termination.
†CI = chain initiation.

are written in terms of the ribonucleotides (containing A, U, C, G; see Chap. 7) in mRNA. The first position refers to the initial nucleotide of the triplet, positioned at its 5′ end. The third nucleotide is at the 3′ end. The term 5′ (five prime) refers to the fact that at this end of the nucleotide chain the last nucleoside does not have a phosphodiester bond on carbon 5 of the ribose. The term 3′ (three prime) refers to the fact that at this end of the nucleotide chain, the last nucleoside does not have a phosphodiester bond on carbon 3 of the ribose. In the interior of the polynucleotide chain the 5′ and 3′ ends of adjacent nucleotides are linked to each other by a phosphodiester bond (see Chap. 7).

Question: What is the codon for methionine?

From Table 17.1, the codon for methionine is seen to be AUG. Methionine is one of two amino acids for which there is only one codon.

Question: What is the significance of the term *CI* in Table 17.1?

This stands for *chain initiation* and indicates that the codon for methionine (AUG) defines the beginning of the translated portion of an mRNA; i.e., methionine is the first amino acid to be incorporated into a polypeptide chain. Less frequently GUG, normally for valine, can function in place of AUG as the chain-initiation codon incorporating methionine (see Table 17.1). It should be noted that while the first amino acid of a newly synthesized polypeptide chain is always methionine (or valine), methionine and valine may also occur within a polypeptide chain.

Question: At what end of a growing polypeptide chain, N or C terminus, is the initiating amino acid found?

It is always incorporated at the N terminus. Thus, the direction of assembly of a chain is $N \rightarrow C$.

Question: Is there an analogous codon that defines the termination site on the mRNA with respect to polypeptide chain formation?

There are actually three codons that function in this way. They are UAA, UGA, and UAG, and are referred to as the *chain termination (CT) triplets* (see Table 17.1), *stop codons*, or *nonsense codons*.

EXAMPLE 17.2

Write a sequence of ribonucleotides that would define a short polypeptide with the amino acid sequence Met-Leu-Arg-Asn-Ala-Val-Glu-Ser-Ile-Cys-Phe-Thr.

A possible sequence is as follows:

5′ AUG UUA CGU AAU GCU GUC GAA UCU AUU UGC UUU ACA UAA 3′

Note the presence of chain initiation and termination codons, respectively, at the beginning and end of the sequence. In translating this sequence, which would occur within a longer mRNA molecule, into a sequence of amino acids, it has been assumed that the codons do not overlap. This has been established experimentally, and for this reason, the triplet genetic code is said to be *nonoverlapping*.

Question: In the ribonucleotide sequence shown in Example 17.2, there is an AUG triplet that can be formed by joining portions of two adjacent triplets, starting from A at the eleventh position. Is it possible for this to function as an alternative initiation codon?

For this to be so, the overall *frame of reading* would be different, i.e., an alternative sequence of triplets would be read to the right of this position. As will be seen below, there are sequences in mRNA, upstream (to the left) of a potential start site, e.g., AUG, that define the frame of reading.

Question: Is the dictionary of amino acid codons shown in Table 17.1 the same for all organisms?

The genetic code dictionary was originally established from studies on the bacterium *E. coli*. It is now known to be the same for all organisms; i.e., it is universal. The only exceptions occur for a few codons in mitochondria from a number of species.

17.3 DNA TRANSCRIPTION IN BACTERIA

As with the establishment of the genetic code and information on the molecular mechanism of DNA replication (Chap. 16), the present detailed knowledge of the mechanism of DNA transcription to produce RNA rests largely upon studies with bacteria, particularly *E. coli*. It is convenient to treat transcription in bacteria first.

Most of the DNA sequences which are transcribed give rise to mRNA, which is subsequently translated into protein. However, the most abundant species of RNA are *ribosomal RNA (rRNA)* and *transfer RNA (tRNA)*, which do not code for protein but function in the process of translation. They are formed by a high level of transcription of a relatively small number of genes (called *rRNA* and *tRNA genes*). In bacteria, transcription of *all* genes is brought about by the enzyme *RNA polymerase*.

Question: What is the nature of the chemical reaction catalyzed by RNA polymerase?

The overall reaction is

$$n\text{NTP} \xrightarrow{Mg^{2+}} (\text{NMP})_n + n\text{PP}_i$$

Thus, it uses the four ribonucleoside triphosphates (ATP, GTP, UTP, and CTP) to assemble an RNA chain, the sequence of which is determined by the template strand of DNA. Nucleotide addition occurs *sequentially*, the phosphodiester bond being formed through the same mechanism as described for DNA polymerase (see Chap. 16, Fig. 16-9). RNA chain growth is in the $5' \rightarrow 3'$ direction. An important distinction between RNA polymerase and DNA polymerase, however, is the ability of the former to start a new chain de novo; i.e., it does not have an obligatory requirement for a primer. The first nucleotide to be incorporated into the chain of RNA contains either adenine or guanine and retains its 5′ triphosphate.

Question: Within a chromosome, a single DNA molecule contains a multitude of genes, and when a particular one is to be transcribed, the RNA polymerase must "know" where to start. How is this achieved?

To transcribe a particular stretch of sequence the RNA polymerase binds to the DNA at a site called a *promoter*, just upstream (i.e., on the 5′ side) of the *transcriptional start site* defined by the template or coding strand.

Many bacterial promoters have been sequenced, and it has been found that, while the sequence of the region is not the same in all cases, there are two segments situated around nucleotide positions 10 and 35 from the start site that vary only slightly from one another, so that a *consensus* sequence can be defined for each. Figure 17-1 shows sequences from a number of promoter regions, which extend through the first or *−10 region* (which is underlined). The −10 region is also called a *Pribnow box* (after the person who discovered it).

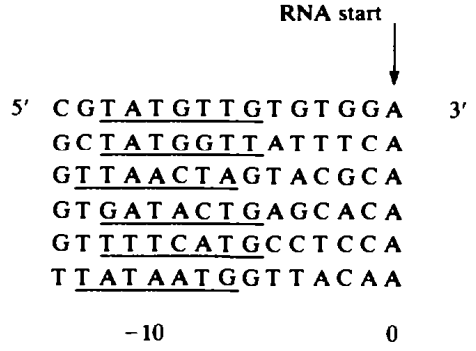


Fig. 17-1 Pribnow box sequences from a number of promoters in bacteria. The regions of homology are underlined.

Another conserved region is centered 35 base pairs (bp) upstream of the start site (i.e., at −35 bp). The consensus sequence is TTGACA: in more detailed form, the sequence is;

$$T_{82}, T_{84}, G_{78}, A_{65}, C_{54}, A_{45}$$

(where the subscript denotes the percent occurrence of the most frequently found base). The distance separating the −35 and −10 sites is important, although the actual sequence in the intervening region is unimportant. It is thought that the distance is critical for locating the RNA polymerase.

Question: How does RNA polymerase recognize the promoter?

Bacterial RNA polymerase is a multisubunit enzyme of molecular weight 480,000. The four major subunits, β, β', α, and σ (M_r = 150,000; 160,000; 36,500; and 86,000, respectively) are present in the ratio 1:1:2:1, and the total complex is more correctly called the *RNA polymerase holoenzyme*.

The σ subunit is involved directly in *promoter recognition*. The complex lacking the σ subunit is called the *RNA polymerase core enzyme*. To start transcription, the σ subunit directs the holoenzyme to a promoter site to form a *binary complex* in which there is a limited unwinding of the DNA duplex to generate an *open promoter complex*. This is the first step in the overall transcription cycle and is called *template binding*.

EXAMPLE 17.3

The transcription cycle is shown in Fig. 17-2.

The template-binding step involves interaction of the holoenzyme, through its σ subunit, with the promoter to give the open promoter complex, as already described. Initiation of the RNA chain can then proceed through the formation of the first phosphodiester bond between ATP (or GTP) and the next nucleotide defined by the template to yield a dinucleoside tetraphosphate:

$$ppp\overset{A}{G} + pppN \rightarrow ppp\overset{A}{G}pN + PP_i$$

Following initiation, the σ subunit is released, and RNA chain elongation proceeds by the sequential addition of nucleotide units to the 3′OH of the previously incorporated nucleotide, as described for the DNA polymerases (Chap. 16). Termination of transcription is effected when the core enzyme reaches a *termination* sequence. In *E. coli* two types of termination sequence have been identified: one requires an additional protein, called *rho*, to effect termination; the other does not. Termination sequences are relatively long (up to ~50 nucleotides) and function through the formation of *hairpins* in the single-strand RNA transcript. Hairpins reflect the presence of *inverted repeat* sequences that allow the RNA chain to bend back on itself and be stabilized through complementary base pairs. Following release of the RNA transcript, the core enzyme is also released. It is then available, after interaction with the σ subunit, for a further cycle of transcription. The RNA released upon termination is called a *primary transcript* because, in some cases, it is modified before being used in a subsequent process.

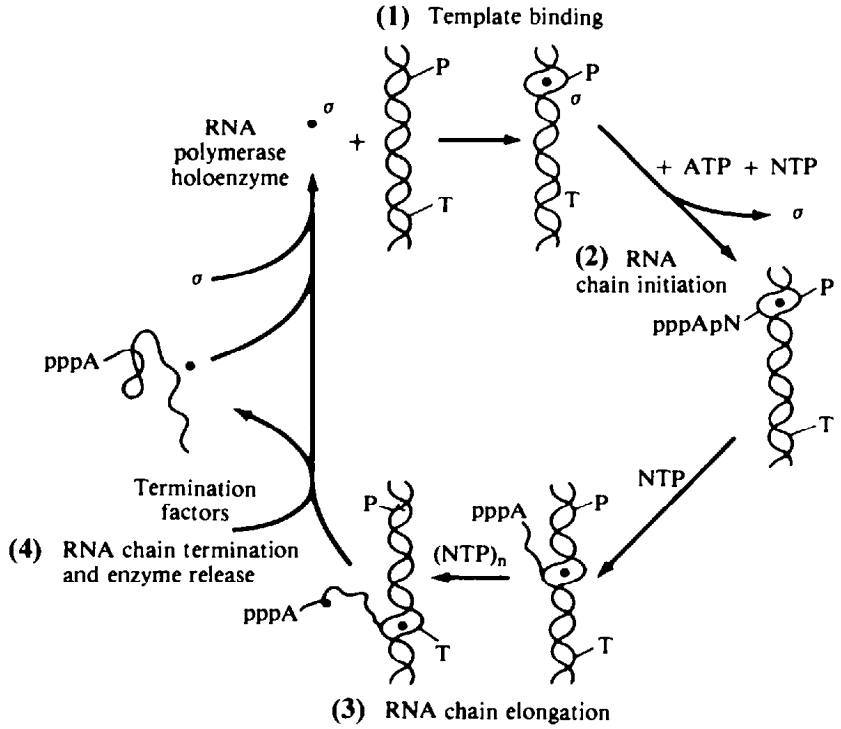


Fig. 17-2 The transcription cycle in bacteria. P and T refer to promoter and termination sites, respectively, for a single RNA transcript. and pppA denotes the triphosphate ATP.

Question: Can a single RNA transcript in bacteria carry information from more than one gene?

Yes. Very frequently groups of adjacent genes are transcribed from a single promoter to give an RNA molecule carrying information for all of them. If this information is to be expressed as proteins, as is normally the case, the single RNA molecule is called *polycistronic mRNA*. (A *cistron* is a genetic unit equivalent to a gene). *Collections* of rRNA and tRNA genes are usually transcribed as single units. In these cases, the primary transcript is modified and subsequently cut by nucleases to yield individual rRNA and tRNA molecules. This is one example of *processing* of primary transcripts, which is much more common in eukaryotes than in prokaryotes.

17.4 DNA TRANSCRIPTION IN EUKARYOTES

The process of transcription in eukaryotes is similar to that in bacteria, but there are important differences.

Question: Is there a single type of RNA polymerase involved in eukaryotic transcription?

No. Eukaryotic RNA polymerases have been isolated from many tissues, and in all cases, three distinct enzymes have been found in the nucleus. All contain a number of polypeptide subunits and are complex in structure. *RNA polymerase I* is known to be involved specifically in the transcription of rRNA genes. *RNA polymerase II* gives rise to transcripts that are subsequently processed to yield mRNA. *RNA polymerase III* is responsible for the transcription of the tRNA genes and a small ribosomal RNA gene that yields a species called *5S RNA*. The three polymerases are distinguishable from one another by their differential sensitivity to the drug *α-amanitin* (the toxic principle of the mushroom *Amanita phalloides*), which does not affect bacterial RNA polymerase. RNA polymerase II is very sensitive to α-amanitin, while RNA polymerase I is completely resistant. RNA polymerase III is moderately sensitive to this inhibitor. Mitochondria have yet another type of RNA polymerase, which is unaffected by α-amanitin but is sensitive to drugs that inhibit bacterial RNA polymerase.

The role of the various subunits of the eukaryotic RNA polymerases has not yet been defined, but presumably there are subunits equivalent to the bacterial σ factor involved in the recognition of promoter sites. As with bacterial promoters, homologies upstream of the start point of transcription have been identified.

Each of the three types of RNA polymerase recognizes a different type of promoter. The promoter for RNA polymerase I is bipartite; i.e., it is made up of *two* conserved regions separated by approximately 70 bp. The RNA polymerase III promoter lies within the transcribed region at around +55 bp. Promoters for RNA polymerase II are extremely complex and diverse. The genes that are transcribed by RNA polymerase II are those that produce proteins and many of these are unique to a particular cell type. Perhaps the most highly (but not universally) conserved region is at −25 bp. This consensus sequence is referred to as the *TATA* or *Hogness box*. Upstream (further along in the 5′ direction on the coding strand) of this there are generally a number of less highly conserved regions which determine when, where and how frequently the gene will be transcribed.

In eukaryotes, regions of DNA as far away as 60 kb from the start site of transcription can influence the level at which the gene is transcribed. These regions have been termed *enhancers*. They can function in either orientation and can lie either at the 5′ or 3′ end of the gene.

17.5 TRANSCRIPTION FACTORS

RNA polymerases require ancillary factors (proteins) for active transcription; these are called *transcription factors*. The enzyme together with these factors constitutes the basal transcription apparatus.

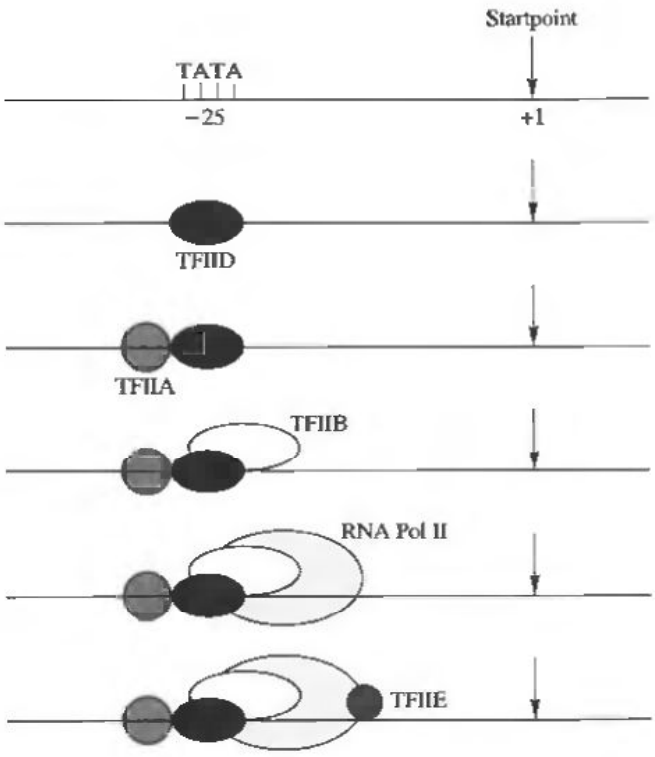


Fig. 17-3 Assembling the basic transcription complex.

For a gene that is transcribed by RNA polymerase II and is expressed in all cell types, the basal factors have been elucidated in some detail. The factors are described as TFIIX, where X identifies the individual factor; e.g., TFIIA, TFIIB, etc. (Fig. 17-3).

The efficiency and specificity with which an RNA polymerase II promoter is recognized depends on short sequences, further away in the 5′ direction than the TATA box, that are recognized by more specialized transcription factors. The positions and combinations of these sites vary from promoter to promoter. Most RNA polymerase II promoters contain six or more of these sites and they lie adjacent to the start site of transcription.

Some of these are fairly common and they bind to transcription factors that are present in all cell types (Table 17.2). These factors in some way alter the ability of RNA polymerase II to transcribe the gene complex.

Table 17.2. Some Common Transcription Factors

Factor	Binding	Distribution
TFIID	TATAA	general
CTF/NF-1	CCAATC	general
SP1	GGGCGG	general
Oct1	ATTTGCAT	general
Oct2	ATTTGCAT	general
NFKB	GGGACTTTC	lymphoid
Erf1	AGATAG	erythroid

Sequencing of enhancer regions has revealed that they contain many of the same conserved sequences (with the exception of the TATA box) that are found in promoters and that bind transcription factors.

Transcription factors must recognize and bind to specific target DNA sequences located in promoters and enhancers. They must also activate transcription. The parts of the protein that are involved in these different functions usually lie in physically independent domains. There are many types of DNA-binding domains. Perhaps the best understood are the zinc fingers, in which a small group of conserved amino acids binds a zinc ion. The activation domains are less well understood. It is thought that they are involved in protein/protein interaction with other transcription factors and/or RNA polymerase.

Question: How is the production of a tissue-specific protein restricted to that one tissue since the gene is present in the nucleus of all cells.

Most of the control of tissue-specific gene expression occurs at the level of transcription; this is achieved with *tissue-specific transcription factors*. For example, all genes that are to be expressed in erythroid cells, e.g., globin, spectrin, and erythropoietin receptor, have the site -AGATA- in their promoters. This site binds to a transcription factor (called *GATA-1*) which is present only in erythroid cells, and the promoter will function only in the presence of this factor.

The main source of information about the sequence requirements of promoters is provided by *mutations*. Mutations in promoters affect the level of expression of the gene they control without altering the gene product.

Question: Are there any human promoter mutants?

Yes, a number of human diseases are caused by point mutations in the promoter regions of important genes. For example, β-thalassemia is a genetic disease in which mutations in the promoter of the β-globin gene result in reduced production of this protein and subsequent anemia. The mutation is usually associated with a reduction in the binding affinity of the promoter for a positive transcription factor.

Question: How can enhancers work on a promoter that is so far away?

It is thought that enhancers function by assembling a complex of transcription factors that interact with the proteins and the promoter in such a way that the intervening DNA is "looped out".

Question: How does histone (Chap. 7) interaction with DNA (to give nucleosomes) affect transcription?

It is generally agreed that the nucleosome structure must be "dissolved" in order for transcription to occur. This dissolution, or removal, of histones must be specific in the case of transcription from restricted regions. The histones interact with the DNA through the positive charge on their basic amino acids, and it has been established that chemical modification to reduce the extent of this charge is one mechanism used to remove histones. *Acetylation* of arginine residues, for example, removes positive charges, while *phosphorylation* and *polyadenosine diphosphate ribosylation* add negative charges. In the last process, the enzyme *poly(ADP-ribose) synthetase* catalyzes the transfer of the ADP-ribose portion of NAD^+ to histones (as well as to nonhistone chromosomal proteins). Exactly how such modifications are effected in response to situations such as hormone stimulation has not been elucidated.

Question: Does transcription in eukaryotes yield polycistronic mRNA?

No, and this is in marked contrast to the situation in bacteria. In eukaryotes, the mRNA, which must be transported out of the nucleus for the purpose of translation, is always *mono*cistronic.

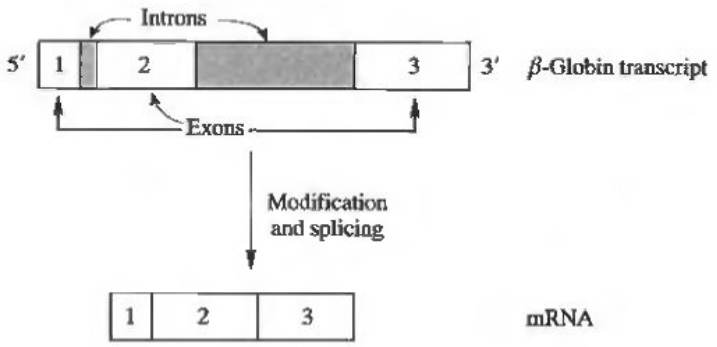


Fig. 17-4 The arrangement of introns and exons in the β-globin gene.

17.6 PROCESSING THE RNA TRANSCRIPT

In bacteria, and prokaryotes in general, the primary transcript provides functional mRNA, ready for translation. In eukaryotes, the vast majority of primary transcripts are chemically modified and have sequences removed from within them before maturing as *functional* mRNAs. This is because eukaryotic genes, which will be expressed as protein, contain nontranslated *intervening* sequences, or *introns*. These are excised, or spliced out, at the primary transcript level to leave what corresponds to the translated segments, or *exons*, in the mRNA. A diagrammatic representation of the β-globin gene is shown in Fig. 17-4.

A primary transcript corresponding to the full length of the gene is first made. This is then chemically modified, and introns (two in the case of β-globin) are removed by splicing. The mixture of primary transcripts present in the nucleus is known as *heterogeneous nuclear RNA* (hnRNA).

Question: What modifications are effected in the primary mRNA transcript?

1. So-called *capping* occurs at the 5′ end of the transcript shortly after its initiation. In the first step of capping, GTP is used to add a guanine nucleotide, which is linked to the chain through a triphosphate bridge; this is catalyzed by *guanyltransferase*. In subsequent reactions, both the added guanine and the first two nucleotides in the primary transcript are methylated. The 5′ cap is thus relatively complex in structure. It has an important role in the subsequent initiation of translation.
2. *Polyadenylation* results in the addition of a *poly(A) tail* of 40–200 residues at the 3′ end of the transcript. The enzyme responsible for this addition is *poly(A) polymerase*. The function of the poly(A) tail is unknown.

The excision of introns must be a very precise reaction if functional messages are to result. Exon-intron junctions within genes have a *consensus sequence*. There are some indications that small RNAs restricted to the nucleus (snRNAs), at least some of which have a sequence complementary to the consensus sequences at the splice junctions, form a secondary structure across two adjacent splice junctions to juxtapose the ends of neighboring exons and provide a framework for cutting and sealing by *processing enzymes*. Splicing of mRNA precursors is carried out by *spliceosomes*, complexes composed of a family of *small nuclear ribonucleoproteins* (snRNPs). (In the case of an rRNA from *Tetrahymena thermophila*, a single intron is spliced out in the presence of only GTP and certain cations.)

EXAMPLE 17.4

The steps involved in the production of a functional mRNA from a typical eukaryotic gene are shown in Fig. 17-5. For this illustration, we assume that the gene has six introns, which are identified as A–F. Exons are represented as 1–7. Not shown in the diagram is the fact that there is probably a defined order for excision of the individual introns.

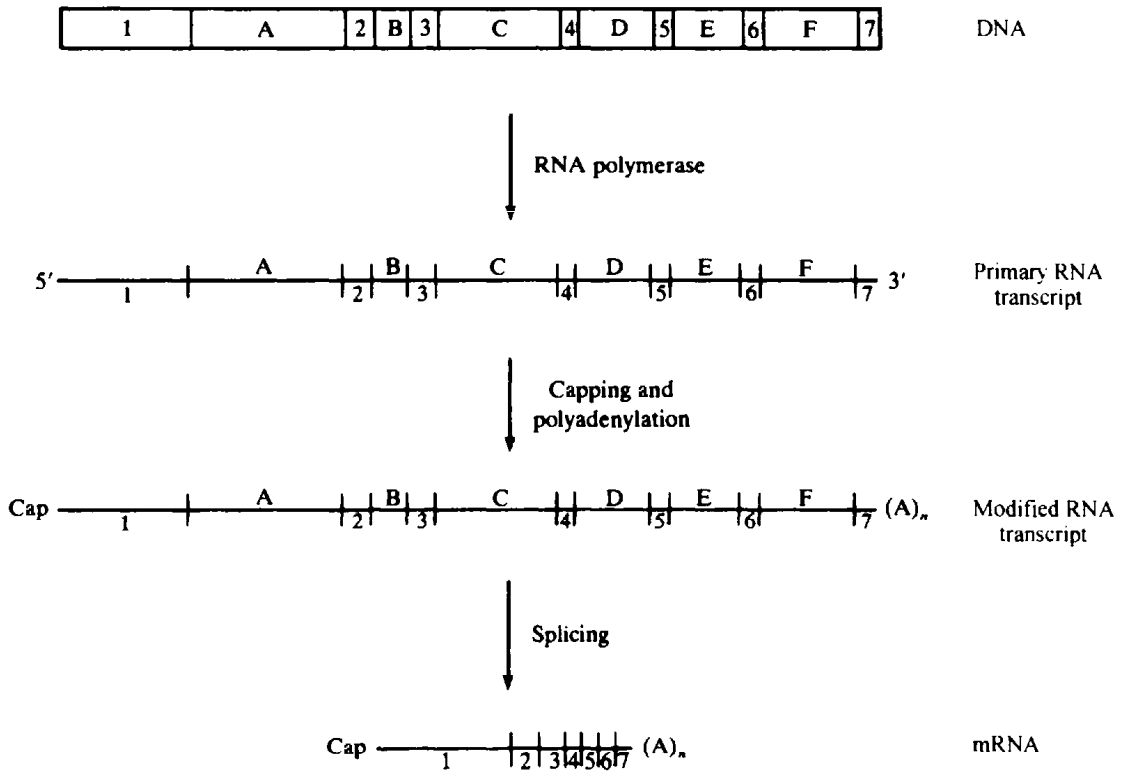


Fig. 17-5 Steps involved in the transcription and processing of a eukaryotic RNA transcript into mRNA.

Question: What is the reason for the presence of introns in most eukaryotic genes?

The precise function of introns is unknown. Nevertheless, there is some evidence that the exons, or coding sequences, give rise to structural domains in the final protein product. It is therefore considered that *interrupted genes* (the exons), as they are called, may reflect an evolutionary process in which combinations of various exons gave rise to different proteins through the joining of different protein *domains*.

17.7 ORGANIZATION OF THE GENOME

Eukaryotic DNA can be divided into three types based on the overall nature of the sequence: highly repetitive, moderately repetitive and nonrepetitive.

Highly repetitive DNA makes up about 30% of the genome and is composed of a frequently repeated simple sequence: e.g., . . . GATCGATCGATC . . . . It is found at *telomeres* and *centromeres* and it is not transcribed.

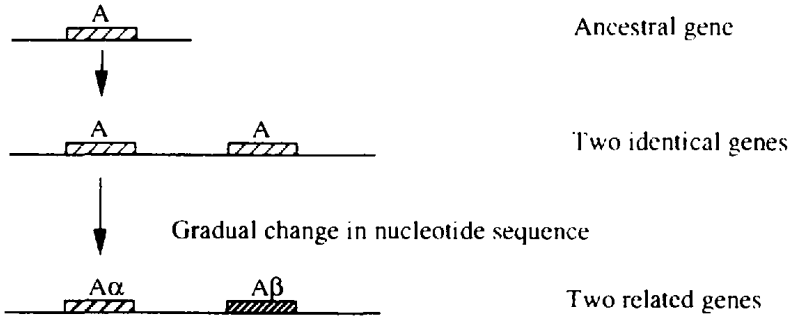
Moderately repetitive DNA makes up about 30% of the genome. Here the DNA sequence itself is more complex and it is repeated less frequently, perhaps 100 times, in the genome. The *ribosomal genes* and the *histone genes* fall into this category.

Nonrepetitive DNA makes up 40% of the genome. Most of this portion is of unknown function, and only about 5% of it is composed of genes which code for proteins.

Question: If we suppose that the average gene is 10,000 bp, then the number of genes in the mammalian genome (3×10^9 bp) would be 300,000. But all the evidence suggests that only 10,000–30,000 genes exist. What is all the rest of the DNA doing there?

We don't really know. It is suspected that this excess DNA may be the result of the way that gene families evolve. Many genes with similar structures lie in clusters that have arisen by gene duplication followed by mutation. This results in the existence of a group of genes with slight differences in sequence which lie side by side. Some members of a gene family may have mutated such that they are no longer functional. These are called *pseudogenes* (see Fig. 17-6).

(a) The human globin gene clusters



(b) A possible scheme for the evolution of a multigene family

Fig. 17-6 The evolution of gene clusters exemplified by the globin genes.

EXAMPLE 17.5

In *E. coli*, at any one time mRNA constitutes only 3–4 percent of the total cell RNA. Only 0.2 percent of the DNA genome is used to code for the 20+ tRNAs, and 0.5 percent for rRNA. Thus, more than 99 percent of the genome serves as a template for mRNA synthesis.

17.8 INHIBITORS OF TRANSCRIPTION

A number of antibiotics function by inhibiting transcription. *Actinomycin D* (see Chap. 16) is an example of one that exerts its effect by binding to the DNA template; it can also block DNA replication.

Question: Would actinomycin D be expected to inhibit transcription in both bacteria and eukaryotes?

Yes, because in binding to the template, it recognizes a structural feature of the DNA double helix and therefore cannot discriminate between the two types of organism.

Question: Are there inhibitors that discriminate between transcription in bacteria and eukaryotes?

Yes. Examples are *rifampicin* and *streptolydigin*, which bind only to bacterial RNA polymerase and block its action; *α-amanitin* binds only to eukaryotic RNA polymerase II and, to a lesser extent, to RNA polymerase III to block their actions.

Question: Bacterial RNA polymerase is a multisubunit enzyme. Are there particular subunits to which rifampicin and streptolydigin bind?

Yes, both bind only to the β subunit. But rifampicin blocks only *initiation* of RNA synthesis, while streptolydigin preferentially blocks *elongation*. This shows that the β subunit is involved in both initiation and elongation of RNA chains.

17.9 THE mRNA TRANSLATION MACHINERY

The sequence of nucleotides in mRNA is converted through the translation "machinery" into a sequence of amino acids that constitutes a polypeptide. This machinery includes tRNA and ribosomes (which contain rRNA and a collection of unique proteins). The function of tRNA is to act as an adapter between the nucleotide sequence (defining the order of codons) and the amino acid sequence to be assembled into a polypeptide.

Question: How does a tRNA molecule function as an adapter between a codon and an amino acid?

The nucleotide sequences of many tRNAs from a wide variety of organisms have been determined. All contain approximately 80 nucleotides, many of them of unusual structure (see Chap. 7). There is at least one tRNA corresponding to each amino acid, and while the sequences within individual tRNAs vary, they all form a common type of secondary structure (cloverleaf) in which the RNA chain folds back on itself to give a maximum amount of base pairing. One part of this structure is involved in the binding of an amino acid, and another part contains a sequence of three nucleotides complementary to one (or more) of the codons for this amino acid. This sequence of three nucleotides interacts with codons in the mRNA during the translational process. There are other features of the structure that are essential for the action of tRNA.

Figure 17-7 is a diagrammatic representation of tRNA folded into the typical cloverleaf structure, containing a number of stems (base-paired) and loops. While the sequences of the different tRNAs are different, there are regions that remain *invariant*. Most of these are in the *loops*, within which the unusual bases are concentrated, and at the 3′ end of the molecule contained within the *acceptor stem*. The sequence at this end is always CCA, and it is to the 3′ OH that the appropriate amino acid is attached through its carboxyl group. The three nucleotides complementary to the codon for the amino acid make up what is known as the *anticodon* (shaded part of Fig. 17-7). The three-dimensional structure of tRNA is known. In this structure, there are additional H bonds, which stabilize the cloverleaf in a more elongated L-shaped structure, with the acceptor sequence at one end and the anticodon loop at the other.

While there is at least one tRNA for each amino acid, there is *not* a separate one for each codon.

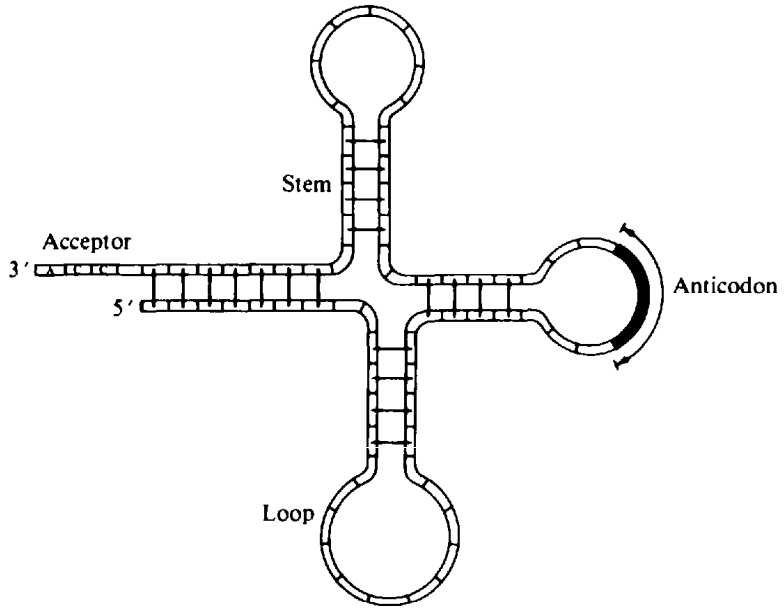


Fig. 17-7 A diagrammatic representation of the folded cloverleaf structure of tRNA.

Question: How can a single tRNA molecule accommodate more than one type of codon?

This can be accounted for by the *wobble hypothesis*: it appears that when a codon in mRNA interacts with the anticodon, unconventional pairing can form between the base in the third position of the codon (3′ end of triplet) and the first position of the anticodon. The unusual nucleoside *inosine* (Chap. 7) frequently occurs in the latter position, and it can pair with A, U, or C. The possibility of more than one type of pairing in this position accounts for the fact that when there is more than one codon for a single amino acid (called *synonyms*, see Table 17.1), the differences are usually in the third position only.

The attachment of an amino acid to an appropriate tRNA is accomplished via *aminoacyl-tRNA synthetase* and the hydrolysis of ATP. There is a separate enzyme specific for each amino acid, and it will recognize all tRNAs for that amino acid. The reaction proceeds in two steps and requires Mg^{2+} (Fig. 17-8). The first step, *amino acid activation*, results in the formation of an aminoacyl-AMP-enzyme intermediate. In the second step, the aminoacyl group is transferred to its appropriate (*cognate*) tRNA, the amino acid being linked to tRNA through an ester bond. It appears that recognition between the synthetase and tRNA is achieved through very precise contact between the

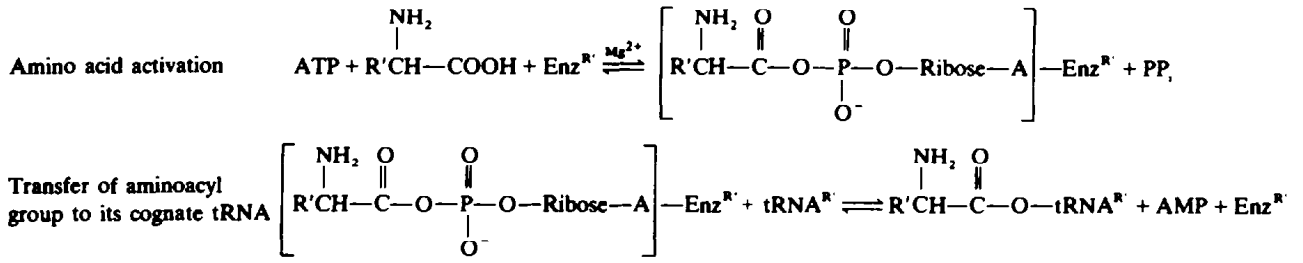


Fig. 17-8 Reactions in the attachment of an amino acid to its cognate tRNA. R′ refers to the amino acid. $Enz^{R'}$ to the appropriate synthetase, and $tRNA^{R'}$ to the cognate tRNA.

two molecules, with single contact points distinguishing one tRNA from another. The first reaction is driven to the right by the hydrolysis of PP_i. Thus, in the overall activation and attachment of an amino acid, two high-energy phosphate bonds are "consumed."

mRNA and aminoacylated tRNAs ("charged" tRNAs) interact on *ribosomes*. The initial interaction occurs in such a way as to allow the codon for the initiating amino acid (methionine) to interact with its appropriately "charged" tRNA to commence polypeptide synthesis.

EXAMPLE 17.6

Ribosomes comprise small and large subunits, distinguishable from one another by their different rates of sedimentation in a centrifuge cell (Chap. 4). The small subunit has a special role in the initiation of polypeptide synthesis. In bacteria, the small and large subunits have sedimentation coefficients of 30S and 50S, respectively. They interact to give a 70S ribosome. In polypeptide synthesis, the interaction occurs as an early step in the overall process. In eukaryotes, the subunit makeup is similar, with other subunits slightly larger. The small (40S) and large (60S) subunits yield an 80S ribosome. The individual subunits in both types of ribosome have the same functions, determined by the types of RNA and proteins present within them. These are listed in Table 17-3.

Table 17.3. Components of Bacterial and Eukaryotic Ribosomes

Bacteria	Eukaryotes
70S ribosome:	*80S ribosome*:
30S subunit	40S subunit
= 16S RNA + 21 proteins	= 18S RNA + ~30 proteins
50S subunit	60S subunit
= 23S RNA + 5S RNA + 34 proteins	= 28S RNA + 5.8S RNA + 5S RNA + ~50 proteins

A considerable amount of information on the precise architecture of the small and large subunits of bacterial ribosomes, defining the surface location of the many proteins, and on the manner of interaction of the subunits is available. Some of this is illustrated in Fig. 17-9; refer to the subsequent text for further explanation of this figure.

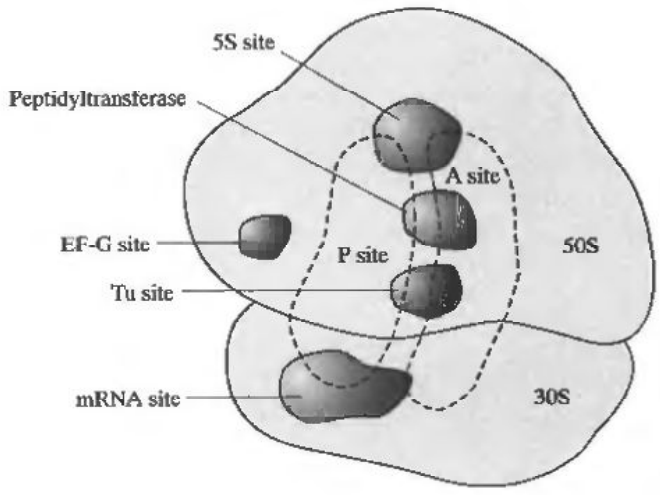


Fig. 17-9 A diagrammatic representation of some of the sites on the *E. coli* ribosome. The peptidyl (P) site accommodates fMet-tRNA$_f^{Met}$, which is involved in initiation of the polypeptide chain. The A site accommodates aminoacyl-tRNA. The protein Tu (*T* for transfer, and *u* for unstable when heated) releases the aminoacyl-tRNA to the A site. EF-G refers to elongation factor G.

17.10 RNA TRANSLATION IN BACTERIA

Because we have a clearer picture of the way in which ribosomes interact with mRNA and assemble polypeptides in bacteria, this will be considered in some detail first. The overall process in eukaryotes is very similar, and the special features of eukaryotes will be treated in the section that follows.

Translation of an RNA message into a polypeptide occurs in three stages: *initiation*, *elongation*, and *termination*. As already mentioned, initiation in bacteria involves the interaction of the 30S ribosomal subunit at the appropriate location on mRNA.

Question: What features of mRNA structure enable interaction with the 30S subunit?

Toward the 5′ end of mRNA, there is a region of 20 or so nucleotides before the initiation codon AUG is reached. This *leader* region contains a sequence responsible for interaction of the mRNA with the 30S subunit. It is known as the *Shine-Dalgarno* (S-D) sequence, and it can bind to a complementary sequence at the 3′ end of the 16S rRNA to position the 30S subunit appropriately for initiation. Other sequences in the leader region are possibly involved in the overall process of initiation of translation, which involves also the binding of the appropriately charged methionyl-tRNA opposite the AUG codon.

Question: Is there any special feature of methionyl-tRNA, in addition to the presence of an anticodon for AUG, that is required for its participation in initiation of translation?

Yes. There are two tRNAs for methionine, distinguishable by their capacity when charged with methionine to be *formylated* by a *transformylase*. The two species are called $tRNA_f^{Met}$ and $tRNA_m^{Met}$. The former is capable of being formylated to yield *N*-formylMet-tRNA$_f^{Met}$ (or fMet-tRNA$_f^{Met}$, for short) and is the species involved exclusively in initiation of the polypeptide chain. Presumably, unique features of the structure of the RNA in this case are required for the initiation process.

In addition to mRNA, fMet-tRNA$_f^{Met}$, and the ribosome subunits, three *initiation factors* (proteins) and GTP are involved in the initiation of polypeptide synthesis. The process is described in Example 17.7.

EXAMPLE 17.7

The first step of RNA translation begins with the initiation of polypeptide synthesis (Fig. 17-10). GTP is bound into the 30S initiation complex and is subsequently hydrolyzed and released upon binding to the 50S subunit. The fMet-tRNA$_f^{Met}$ occupies what is known as the peptidyl (P) site of the ribosome (Fig. 17-9); another site (A), capable of accommodating an aminoacyl-tRNA, is empty at this stage. It is aligned with the next codon (shown as xxx in Fig. 17-10) in the mRNA.

Question: Is there any special mechanism for filling the A site with the appropriate aminoacyl-tRNA?

Yes. Transport of the appropriately charged tRNA to the A site requires association with a protein-GTP complex. The protein, called *Tu* (so-named because it is a *transfer* factor and is *unstable* when heated) is known as an *elongation factor*. Upon releasing the aminoacyl-tRNA to the A site, the Tu-GTP is hydrolyzed to Tu-GDP + P_i. Tu-GTP is then regenerated from Tu-GDP through reactions involving another protein, *Ts* (a heat-*stable transfer* factor), and GTP. Thus, one high-energy phosphate bond is consumed in binding the aminoacyl-tRNA into the A site. Everything is now ready for peptide bond formation (elongation).

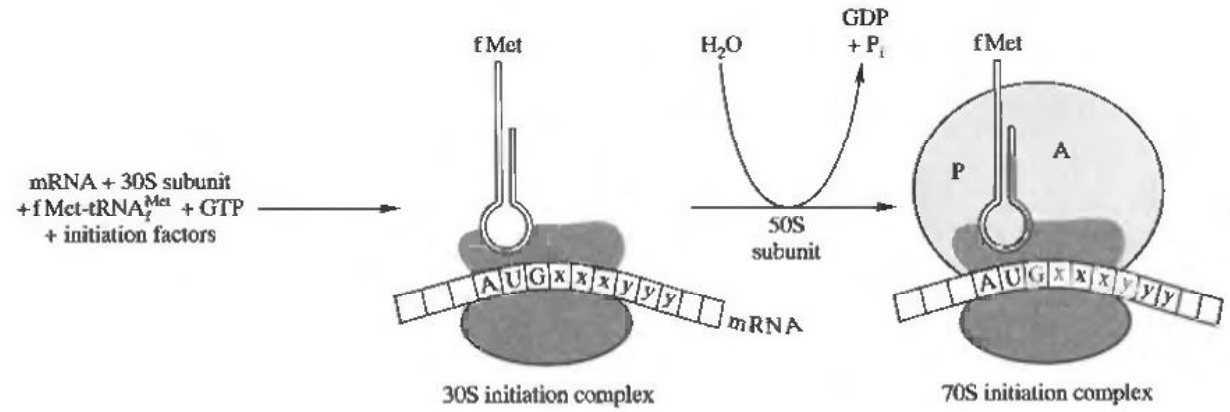


Fig. 17-10 Initiation of polypeptide synthesis.

The elongation phase of polypeptide synthesis and its termination are described in Example 17.8. The A site is shown to be filled by AA_2-tRNA where the codon xxx is located.

EXAMPLE 17.8

The second step of RNA translation involves elongation of the polypeptide chain (Fig. 17-11). One of the protein components of the 50S subunit is a *peptidyltransferase*. As the name implies, it transfers the fMet (and in later reactions, a peptide) from the P site to the A site. To do this, the ester bond linking fMet to its tRNA is broken and the aminoacyl is transferred to the amino group of the adjacent aminoacyl-tRNA (AA_2-tRNA in Fig. 17-11) to form the first peptide bond. In the next step, a *translocase* called *elongation factor G* (EF-G), in association with GTP hydrolysis, shifts (or translocates) the ribosome by one codon to position the dipeptidyl-tRNA in the P site, leaving the A site available for the binding of another aminoacyl-tRNA. This process of aminoacyl-tRNA binding, peptide bond formation, and translocation continues until a *stop codon*, which defines the completion of the polypeptide chain, is aligned with the empty A site.

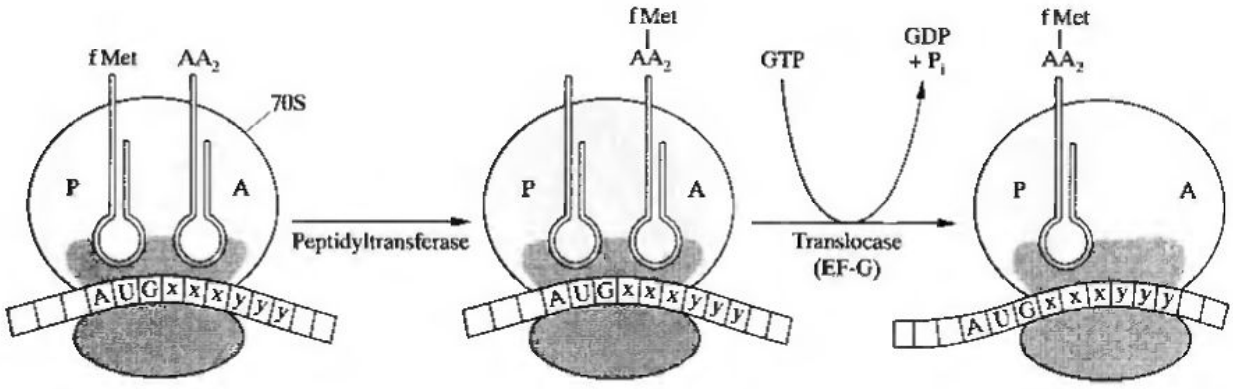


Fig. 17-11 Elongation step of polypeptide synthesis.

Question: How is the long polypeptide chain released from its ester linkage to the tRNA in the P site?

In bacteria, there are three *release-factor proteins*: RF1, RF2, and RF3. In response to the stop codons, they presumably bind (in various combinations) to the A site and cause hydrolysis of the ester bond to release the chain; in so doing, they generate the free carboxyl terminus of the polypeptide.

Question: During translation, as the ribosome moves along the mRNA, it leaves the leader region (containing the *ribosome binding site*) empty. Is it possible for initiation of a new polypeptide chain to occur before completion of a previous one?

Yes. It is common for any single mRNA to be translated simultaneously by many ribosomes. They give rise to a structure called a *polyribosome* or *polysome*.

17.11 RNA TRANSLATION IN EUKARYOTES

The molecular mechanism of translation in eukaryotes is very similar to that in bacteria. The activation of amino acids and attachment to tRNAs and the steps of initiation, elongation, and termination of polypeptide chains are essentially the same in overall terms. The small and large ribosomal subunits of bacteria and eukaryotes are equivalent with respect to their roles in initiation and elongation of chains.

There are two *significant* differences between the mechanism of translation in eukaryotes and in bacteria, and they relate to initiation of translation. First, while there are two forms of tRNA for methionine in eukaryotes, one of which is used in initiation, neither charged form is formylated; the transformylase is not present in eukaryotes. (Eukaryotic Met-$tRNA_i^{Met}$, however, can be formylated by the bacterial enzyme.) A second difference, of more significance, is the involvement of the methylated 5′ cap of the mRNA in initiation of translation. If the cap is missing, translation is inefficient. It has been established that binding of the 40S ribosomal subunit to the leader region of mRNA requires additional factors called *cap-binding proteins*. There is evidence to suggest that the cap is the major structural feature needed for 40S subunit binding.

17.12 POSTTRANSLATIONAL MODIFICATION OF PROTEINS

Most polypeptides synthesized on ribosomes are later chemically modified. Thus the formyl group on the N-terminal methionine in polypeptides of bacteria is removed by a *deformylase*. In both bacteria and eukaryotes, the N-terminal methionine, sometimes along with a few additional amino acids, is removed by *aminopeptidases*.

Question: The amino acids *hydroxyproline* and *hydroxylysine* are absent from Table 17.1, which describes the genetic code. How do these amino acids arise in some proteins?

Hydroxyproline and hydroxylysine occur most noticeably in collagen. These are formed by modification of proline and lysine residues by specific enzymes *after* synthesis of the collagen chains. It is interesting to note that *prolylhydroxylase*, which hydroxylates proline, requires *ascorbate* (vitamin C) as a coreactant. Other chemical modifications known to occur commonly are the attachment of sugars (glycosylation) to asparagine, serine, and threonine residues and the phosphorylation of serine. Chemical modifications are also associated with the transport of proteins out of the cells in which they are synthesized.

Question: How can proteins be transported through a hydrophobic membrane and out of a cell?

Proteins destined for *secretion* from cells (in bacteria and eukaryotes) are usually synthesized in a *precursor form*. This form contains what is called a *signal sequence*, which is relatively hydrophobic

and consists of 15–30 amino acid residues, at the N terminus. This sequence, as it is formed on the ribosome, somehow attaches to a membrane and penetrates it. As the polypeptide chain continues to elongate, it passes through to the other side of the membrane. In the meantime, the signal sequence is removed by a *signal peptidase*. In the case of bacterial cells, such a process can lead to transport of the protein into the *periplasmic space* (between the inner and outer membranes at the surface) or out of the cell altogether. In the case of eukaryotes, the signal peptide enables the transfer of the protein into the lumen of the endoplasmic reticulum. From here it is transported by other mechanisms out of the cell. Insulin provides a good example of the transport or secretion of a protein out of a eukaryotic cell.

EXAMPLE 17.9

Insulin is formed in special cells of the pancreas (the β cells of the islets of Langerhans). The immediate product of mRNA translation is a single polypeptide called *preproinsulin*. The modifications associated with the conversion of preproinsulin to insulin are shown in Fig. 17-12.

Porcine preproinsulin is a single polypeptide of 107 amino acid residues. After the signal peptide is synthesized on the ribosome, it recognizes a receptor on the endoplasmic reticulum and attaches to it. (The attachment of ribosomes to the endoplasmic reticulum through peptide chains is responsible for what is known as the *rough endoplasmic reticulum*, Chap. 1.) The polypeptide chain passes through the membrane and into the lumen, where the signal peptide of 23 residues is cleaved off to yield a shortened chain of 84 residues; the latter folds on itself to form intramolecular disulfide bridges joining cysteine residues. This folded molecule is called *proinsulin*.

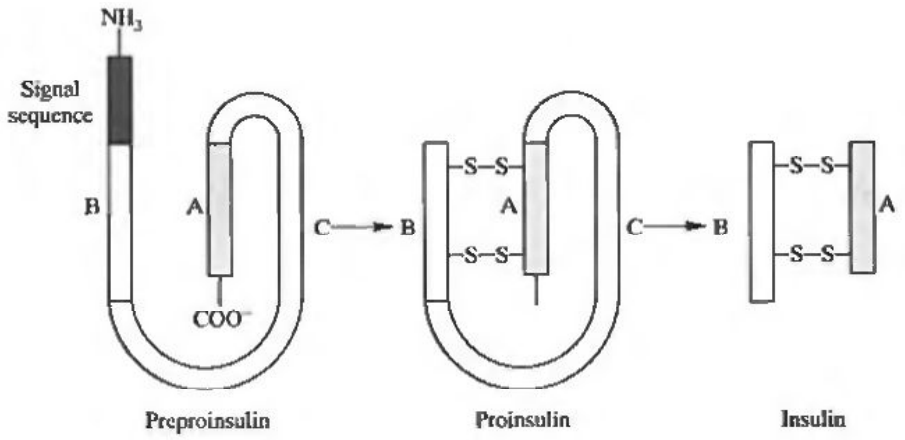


Fig. 17-12 Conversion of preproinsulin into insulin.

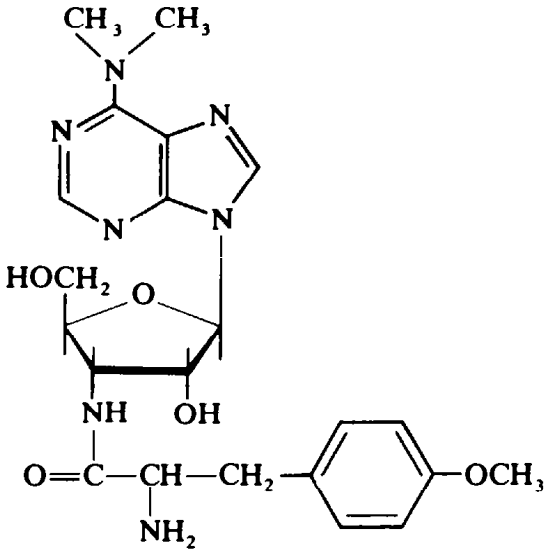
Proinsulin is packaged into membrane vesicles and transported to the *Golgi apparatus*, where conversion to insulin begins. This includes cleavage of the uncross-linked chain (C), thus removing the "connecting" peptide. Conversion is completed as the molecule is transported via *secretory granules* from the Golgi to the *cytoplasmic membrane*, with which the granules fuse to release the mature insulin, now consisting of two disulfide-cross-linked polypeptide chains, into the blood stream.

17.13 INHIBITORS OF TRANSLATION

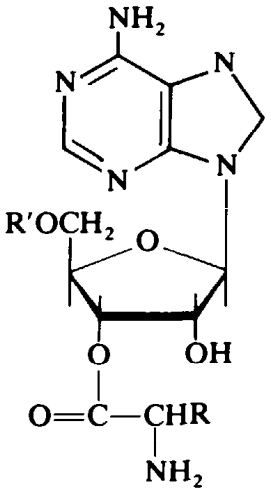
Because of the large number of steps associated with the translation of mRNA into protein, there are numerous opportunities available for blocking it with inhibitors. The action of many antibiotics is based on blocking translation in bacteria.

Question: Which inhibitors will block translation in both bacteria and eukaryotes?

Good examples are *fusidic acid* and *puromycin*. The former inhibits the binding of charged tRNA to the A site of the ribosome. Puromycin acts by virtue of its similarity in structure to an aminoacyl-tRNA (see below).



Puromycin

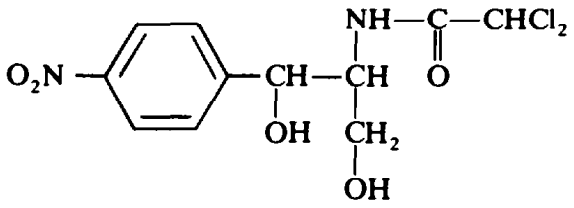


Aminoacyl-tRNA

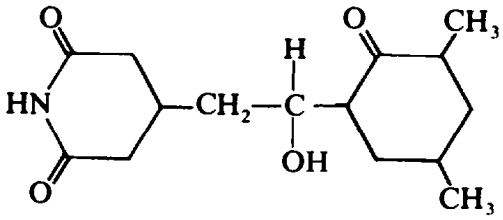
It competes with the latter as an acceptor in the peptidyl transfer reaction. The growing chain is transferred to the NH_2 group of puromycin and is prematurely terminated.

Question: Are there different inhibitors that act on analogous targets in bacteria and eukaryotes?

Yes. *Chloramphenicol* inhibits the peptidyltransferase of the 50S ribosomal subunit of bacteria, while *cycloheximide* inhibits the analogous enzyme in the 60S subunit of eukaryotic ribosomes. Their structures are shown below.



Chloramphenicol



Cycloheximide

It is interesting to note that chloramphenicol blocks translation in mitochondria, further indicating the similarity of mechanisms of gene expression in mitochondria and bacteria.

There are many other inhibitors specific for just bacterial or eukaryotic cells. An interesting one is *diphtheria toxin*, which is effective only in eukaryotes.

EXAMPLE 17.10

Diphtheria toxin, produced by *Corynebacterium diphtheriae*, is a single polypeptide (M_r = 63,000) with two intrachain disulfide bridges. A portion of the molecule, the A fragment (M_r = 21,000), must enter the cytoplasm of a cell to exert the toxic effect. Through a fairly complex mechanism, the larger molecule is cleaved into two on the membrane surface to yield A and B fragments. The latter facilitates entry of the A fragment into the

cytoplasm of the cell, where it specifically catalyzes the chemical modification (ADP-ribosylation) of the translocase (called EF2) to *inactivate* it and block polypeptide chain growth.

17.14 CONTROL OF GENE EXPRESSION

The end products of gene expression are proteins, mainly enzymes, and it is essential that their levels be strictly controlled. There are many potential sites of control in both bacteria and eukaryotes. DNA or gene amplification (Chap. 16) in eukaryotes is one way of responding to the demand for more of the protein product; if there are more copies of the gene, then transcription can occur at a faster rate. More often, control is effected at the level of either transcription or translation, with the former probably being more important for both bacteria and eukaryotes. Transcriptional control in bacteria is particularly effective because of the very short half-life (a few minutes) of mRNA in such cells; the half-life is longer in eukaryotes. The prototype for transcriptional control is the *lactose operon* in *E. coli*.

EXAMPLE 17.11

The lactose operon (or *lac* operon), a region of ~5.3 kb of the *E. coli* chromosome, contains the genes coding for the enzymes responsible for lactose metabolism. The genes for the three enzymes involved, *β-galactosidase*, *galactoside permease*, and *thiogalactoside transacetylase*, are situated next to one another in this segment of DNA. They are transcribed as a single unit of RNA, with transcription being controlled by sequences (control elements) towards the 5′ end of the 5.3-kb segment of DNA. The control elements comprise a *promoter* (to which RNA polymerase binds) and an adjacent *operator* (to which a *repressor protein* can bind to block transcription by the RNA polymerase). The repressor is produced by a gene located on the 5′ side of the promoter. This is illustrated in Fig. 17-13. Because all three genes, *z*, *y*, and *a*, are transcribed as a single unit (polycistronic mRNA), they are said to be expressed *coordinately*. When transcription is blocked by the repressor, none of the genes are expressed.

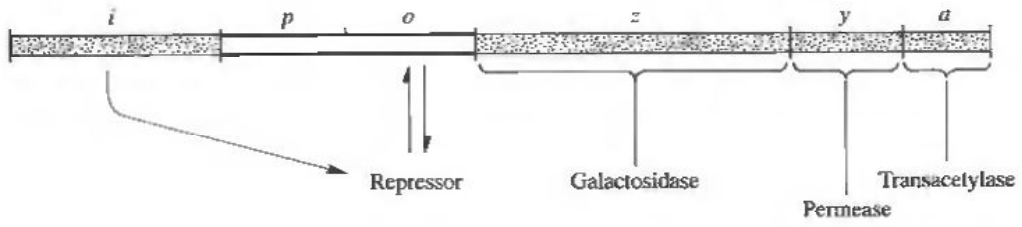


Fig. 17-13 The lactose operon of *E. coli*. Abbreviations: i = repressor gene; p = promoter; o = operator; z = β-galactosidase gene; y = permease gene; and a = transacetylase gene. The i gene product is the repressor.

Question: How does the cell overcome repression to express the genes of the *lac* operon?

When lactose is added to a cell culture, a small amount of it enters the cell and is converted to *allolactose*. This metabolite binds to the repressor, altering its conformation and causing it to be released from the operator. RNA polymerase is then free to transcribe the three genes, and the allolactose is said to be acting as an *inducer*. This relief of repression by an inducer is an example of *negative control* of expression.

There is also an aspect of *positive control* in the lac operon. The *catabolite activator protein* (CAP), carrying bound cAMP, is required for the binding of RNA polymerase to the promoter; i.e., it has a direct, positive effect on transcription. However, relief of repression (i.e., induction) will not occur in the presence of glucose, because glucose lowers the level of cAMP, so that CAP is unable to exert its effect. This reflects the preference of the cell to use glucose rather than lactose as a carbon source. Thus it can be seen that the cell stringently controls expression of the *lac* genes; it expresses them only if it needs to metabolize lactose.

Question: Are there good examples of transcriptional control in eukaryotes?

It has been established that many of the *steroid hormones* (Chap. 13) act by stimulating transcription. Through a series of complex steps, involving interaction with a cytoplasmic receptor, the hormone enters the nucleus, where it binds, either by itself or in association with the receptor or other proteins, to a specific site on the DNA. Exactly how it then induces transcription is still not clear, but both positive and negative control mechanisms seem to be involved. A well-studied example is the estrogen-mediated induction of ovalbumin mRNA synthesis in the chicken oviduct.

Eukaryotes potentially have many more opportunities for control of gene expression than do bacteria. For example, the cell could take advantage of control at the level of the *processing of primary transcripts*. It is known that RNA is not transported across the nuclear membrane until all introns are excised. A more subtle form of control could involve *alternative modes of splicing* a particular transcript. There are now examples known where this occurs to yield different mRNA molecules. Perhaps one of the best-known examples of yet another level of control in eukaryotes is that of translational control of *globin* synthesis.

EXAMPLE 17.12

Globin is synthesized in reticulocytes (see Chap. 1, Prob. 1.1), which have no nucleus and therefore cannot utilize transcriptional and other potential modes of control. Control of globin synthesis from the pool of globin-enriched mRNA is geared to the concentration of *hemin* [Fe(III)-protoporphyrin], which has the ability to inactivate a translational inhibitor of protein synthesis. The inhibitor is a *protein kinase* that phosphorylates and inactivates one of the initiation factors involved in initiation of translation. When the concentration of hemin is high, it binds to a regulatory subunit of the kinase and, as a result, initiation of globin synthesis can proceed.

However, by far the most important form of regulation of gene expression in eukaryotes occurs at the level of transcription. The human genome is 1000 times larger than the *E. coli* genome but the number of genes is only about 50 times greater. The proteins are probably on average no larger. So what is all the extra DNA doing? It is possible that most of the eukaryotic genome is involved in gene regulation, i.e., the decision to transcribe or not. Enhancer elements, which clearly play a role in this process (see Section 17.4) are often found to lie large distances from the coding sequences of many genes, and in some cases as far away as 20 kb. Some elements called *locus control regions* (*LCRs*), many of which also function as enhancers, have been shown to maintain large regions of chromatin in an "open" structure. "Open" chromatin is classified as *euchromatic* as opposed to *heterochromatic*. Heterochromatic chromatin is generally made up of repetitive DNA and it is not transcriptionally active (see Chap. 7). These LCRs seem to be tissue specific, i.e., they function only in the cell type in which the adjacent gene is active. The maintenance of particular regions of chromatin in an "open" configuration in particular cell types may be the major means by which gene expression is regulated in eukaryotes.

Solved Problems

THE GENETIC CODE

17.1. Assuming that the first three nucleotides define the first codon, what is the peptide sequence coded for by (*a*) UAAUAGUGAUAA, (*b*) UUAUUGCUUCUCCUACUG?

SOLUTION

(*a*) None. These four codons are not translatable; they are signals for chain termination.

(*b*) $(Leu)_6$. This illustrates the degeneracy of the genetic code for leucine; there are six codons for leucine, the most for any of the amino acids.

17.2. From within the sequence AAUUAUGUUUCCAUGUCCACCU, identify two possible sites from which initiation of translation could commence and write the sequence of the first three amino acids.

SOLUTION

AUG is the most common initiation codon. The first one encountered in the sequence would give the following sequence for nine nucleotides from this position: AUGUUUCCA. This codes for Met-Phe-Pro. Farther along there is the sequence AUGUCCACC, which codes for Met-Ser-Thr.

17.3. Within the coding strand of the DNA in a gene, the codon sequence ATA is changed by mutation to ATG. Following replication of the DNA, what change in the polypeptide product would this cause?

SOLUTION

The coding strand of DNA gives the same sequence as the mRNA, except for T being present in the DNA in place of U. Thus, the codon will change from AUA to AUG. This causes isoleucine to be replaced by methionine.

17.4. An mRNA contains the following translated sequence, with the reading frame defined by the grouping of triplets:

AUG.CUC.ACU.UCA.GGG.AGA.AGC

(*a*) What amino acids would result from this sequence? (*b*) If the first C (nucleotide containing cytosine) encountered is deleted from the sequence, what new amino acid sequence would result?

SOLUTION

(*a*) The original sequence would give Met-Leu-Thr-Ser-Gly-Arg-Ser.

(*b*) The new sequence would be

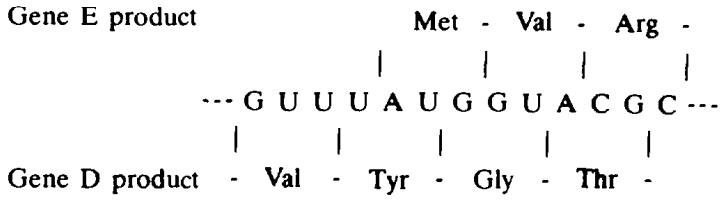
AUG.UCA.CUU.CAG.GGA.GAA.GC

which would give Met-Ser-Leu-Gln-Gly-Glu.

17.5. The DNA of the virus ϕX174 appears to contain insufficient nucleotide residues to code for the nine different proteins that are its gene products. How might this arise?

SOLUTION

In this case, the same sequence of nucleotides codes for more than one protein, by the use of more than one reading frame. That is, the sequence of codons for one protein is out of phase with the sequence of the overlapping gene. See the figure below.



17.6. The fidelity of DNA replication is enhanced by a "proofreading" function, whereby errors in the complementary sequence are excised and repaired (Chap. 16). Why is a similar mechanism not found in protein synthesis?

SOLUTION

The consequences of errors in protein synthesis are not as serious. A single defective protein molecule will, in general, not cause deleterious effects; such a protein may not function properly or may be unstable, and may represent an energy wastage to the cell; however, such errors do not become perpetuated in future generations.

17.7. The template strand of a double-helical segment of DNA contains the sequence:

5′ GCTACGGTAGCGCAA 3′

(*a*) What sequence of mRNA can be transcribed from this strand?

(*b*) What amino acid sequence would be coded for, assuming that the entire transcript could be translated?

SOLUTION

(*a*) The transcribed RNA would be complementary to the above strand, and would have U replacing T:

3′ CGAUGCCAUCGCGUU 5′

Written in the 5′→3′ direction this becomes:

5′ UUGCGCUACCGUAGC 3′

(*b*) Translation occurs in the 5′→3′ direction, resulting in the following sequence:

-Leu-Arg-Tyr-Arg-Ser-

DNA TRANSCRIPTION IN BACTERIA

17.8. Would it be possible for a single strand of DNA to function both as a coding and template strand with regard to transcription?

SOLUTION

Yes. For this to occur, transcription by RNA polymerase would proceed through the segment of duplex DNA in *opposite directions* and from different promoters. Each strand of the duplex DNA would be functioning as both coding and template strands, but for two different RNA transcripts.

17.9. Which of the following could not represent the 5′ end of an RNA transcript?

(*a*) pppApGpCpU . . . . .

(*b*) pppUpCpGpA . . . . .

(*c*) pppCpApGpA . . . . .

SOLUTION

Only (*a*) is possible, because in initiating transcription, RNA polymerase always incorporates a pppA or a pppG in the first position.

17.10. Is the Pribnow box the same as a bacterial promoter?

SOLUTION

No. The Pribnow box is only *part* of the sequence that defines the promoter. It corresponds to the -10 bp region of ~7 nucleotides. There is another region centered around the -35 bp position that is an essential part of the promoter.

17.11. Which subunits of bacterial RNA polymerase are needed for initiation of transcription from a promoter?

SOLUTION

First, the σ subunit is needed for promoter binding and the formation of an open promoter complex. Subsequent to this, the β subunit (which will bind the inhibitor rifampicin) is essential for the formation of the first phosphodiester bond.

17.12. Is the transcription termination sequence incorporated into the RNA transcript?

SOLUTION

Yes. After the RNA polymerase transcribes the inverted repeat sequence, it is the ability of this region to form a hairpin in the single-strand transcript that is responsible, at least in part, for termination of the chain.

17.13. How many transcription termination sequences would be present in polycistronic mRNA?

SOLUTION

Polycistronic mRNA, which forms as a primary transcript in bacteria, is a continuous length of RNA transcribed from a single promoter. It will therefore contain only one *normal* termination sequence (i.e., ignoring a possible *attenuator* sequence before the first initiation codon).

17.14. Does processing of RNA transcripts occur in bacteria?

SOLUTION

Processing refers to modification of the primary transcripts formed by RNA polymerase. In bacteria, it is restricted to the transcripts that contain rRNA and tRNA. In these cases, larger transcripts are chemically modified and then cut down to the smaller, mature forms of rRNA and tRNA by nucleases.

DNA TRANSCRIPTION IN EUKARYOTES

17.15. How many types of RNA polymerase might you expect to find in a eukaryotic cell?

SOLUTION

Four. In the nucleus, there are the RNA polymerases I, II, and III. In mitochondria, there is another RNA polymerase, which is similar to bacterial RNA polymerase.

17.16. With respect to the general mechanism of RNA chain growth, are there any differences among the various types of RNA polymerase in eukaryotic cells?

SOLUTION

No. All RNA polymerases use duplex DNA as a template and copy one of the strands; they synthesize RNA in the $5' \rightarrow 3'$ direction and use ribonucleoside triphosphates as substrates.

17.17. Would the major transcriptional activity of a eukaryotic cell be affected by α-amanitin?

SOLUTION

No. The most abundant RNA component of a cell is always rRNA (Chap. 7). The rRNA genes are transcribed by RNA polymerase I, which is resistant to α-amanitin.

17.18. Is it possible that α-amanitin exerts its inhibitory effect on certain eukaryotic RNA polymerases by interfering with the availability of the substrates?

SOLUTION

No. All RNA polymerases use the same substrates, the α-amanitin inhibits some RNA polymerases but not others.

17.19. Does monocistronic mRNA of eukaryotes generally represent a primary RNA transcript?

SOLUTION

No. In the vast majority of cases, the mRNA, which is always monocistronic in eukaryotes, is formed by the processing (modification and splicing) of primary transcripts.

INHIBITORS OF TRANSCRIPTION

17.20. Would actinomycin D be expected to block transcription of rRNA and tRNA genes and those that code for protein products?

SOLUTION

Yes. Actinomycin D blocks transcription by binding to the DNA template. In doing so, it recognizes a common structural feature of all duplex DNAs, binding by intercalation between stacked base pairs (Chap. 16).

17.21. The bacterial RNA polymerase inhibitors rifampicin and streptolydigin each bind to the same subunit of the enzyme, but their overall effect on the enzyme's activity is different. Why is this so?

SOLUTION

Each of these inhibitors binds exclusively to the β subunit of RNA polymerase. This subunit is involved in both initiation and elongation of RNA chain growth. Rifampicin must bind to the subunit in a way that affects only the initiation step; it has no effect on elongation. Streptolydigin, on the other hand, binds in a manner that blocks both activities.

THE mRNA TRANSLATION MACHINERY

17.22. How many varieties of rRNA and tRNA species are present in a cell actively involved in translation?

SOLUTION

At the most, there are 3–4 types of rRNA, depending on whether they are bacterial or eukaryotic. On the other hand, there will be *at least* 20 types of tRNA (at least one for each amino acid) in any type of cell.

17.23. Why is it necessary for all tRNAs to have similar overall dimensions?

SOLUTION

All tRNAs must have similar overall dimensions because, during translation, all charged tRNAs interact singly and very precisely with the same sites on the ribosome. The anticodon is positioned at one "end" to allow it to interact with the bound mRNA, and the amino acid is precisely located on the surface of the ribosome with respect to the location of the bound peptidyltransferase.

17.24. Would it be possible for a single tRNA to accommodate all the codons for leucine?

SOLUTION

Leucine has six codons, the largest number for any amino acid. According to the *wobble hypothesis*, a single tRNA can accommodate more than one codon. The wobble hypothesis allows for up to three different nucleotides, but only at the third position in the codon, to interact with a single nucleotide in the anticodon. The fact that there are six codons in leucine means that they must differ at positions other than the third, and they therefore could not be accommodated by a single anticodon in a particular tRNA molecule.

17.25. Aminoacyl-tRNA synthetase functions in two steps to bring about the attachment of an amino acid to its cognate tRNA. Does it show specificity for either the amino acid or the tRNA at each of these steps?

SOLUTION

In the first step (amino acid activation), the enzyme recognizes its appropriate amino acid. In the second step, the amino acid must be attached to the correct tRNA, and it is therefore essential that the latter be recognized by the enzyme. (It does happen that sometimes, at a low frequency, a similar but incorrect amino acid is incorporated in the first step: the incorrect amino acid is released when the enzyme recognizes its appropriate tRNA.)

RNA TRANSLATION IN BACTERIA

17.26. What is the "leader region" of bacterial mRNA?

SOLUTION

The leader is the region between the 5′ end of the RNA and the initiation codon. It contains untranslated sequences that are involved in ribosome binding, a step essential for translation of the message into a polypeptide.

17.27. Why can't $tRNA_m^{Met}$ function in initiation of polypeptide synthesis from an appropriate AUG codon?

SOLUTION

In initiation of translation, the methionyl-tRNA species $tRNA_f^{Met}$ is used exclusively. Presumably, the unique nucleotide sequence of this species is required for the initial interaction with the small ribosomal subunit, the leader region of mRNA, initiation factors, and GTP to give the first initiation complex. It is unlikely that formylation (which can occur only with methionyl-$tRNA_f^{Met}$) is essential for initiation, as this is not the case in eukaryotes. It thus appears that $tRNA_m^{Met}$ is excluded from initiation because essential structural requirements in the RNA are not met.

17.28. How many ATP *equivalents* are consumed with the incorporation of an amino acid into a polypeptide?

SOLUTION

One ATP is used for charging of the tRNA, and then one GTP at *each* of the steps of binding aminoacyl-tRNA to the A site of the ribosome, and translocation. Thus, ignoring initiation, the equivalent of three ATPs are used for each amino acid incorporated. But remember that in amino acid activation, the products are AMP and PP_i, the latter being hydrolyzed to P_i to drive the reaction to completion. Thus, the equivalent of four high-energy phosphate bonds are used for each amino acid incorporated.

17.29. In translocation, the peptidyl-tRNA is shifted from the A site to the P site on the ribosome. What happens to the peptidyl-tRNA anticodon-codon interaction?

SOLUTION

This must remain undisturbed, and it is probably clearer to say that the whole ribosome moves by one codon position with respect to the peptidyl-tRNA–mRNA complex.

17.30. What major structural requirement for initiation of translation of eukaryotic mRNA does not exist for bacterial mRNA?

SOLUTION

During the processing of primary transcripts destined for mRNA formation in eukaryotes, a methylated guanine "cap" is attached to the 5′ end. This is essential for efficient initiation of translation of the mature eukaryotic mRNA and does not occur in prokaryotes.

17.31. Is there any obvious reason why polysomes could not form in eukaryotes?

SOLUTION

Polysomes form as the result of the loading of sequential ribosomes at the 5′ end of mRNA such that many ribosomes are progressing through a single mRNA at any one time. Clearly, polysomes do form in eukaryotes. Once the first ribosome has moved a significant distance from the ribosome binding site, there is no reason why a subsequent one should not be loaded through the normal sequence of events.

17.32. *Coupled transcription-translation* in prokaryotes refers to the commencement of translation of an RNA molecule before its transcription from the DNA template is complete. Could such a situation arise in eukaryotes?

SOLUTION

No. Transcription in eukaryotes must be completed within the nucleus. The mRNA finally produced is transported to the cytoplasm for translation into a polypeptide.

POSTTRANSLATIONAL MODIFICATION OF PROTEINS

17.33. When proteins are hydrolyzed with acid at high temperature, they are broken down to their constituent amino acids. Commonly, the amino acid cystine is found among them. But cystine is not included among the amino acids listed in the dictionary of codons. Why?

SOLUTION

Cystine is composed of two molecules of cysteine linked through oxidation of their —SH groups to give a disulfide bond. Such oxidation, which is important in stabilizing the folded structure of some proteins, represents a posttranslational modification of a protein. Thus, cystine is never incorporated as such into a polypeptide during translation, and there is no codon that corresponds to it.

17.34. The signal sequence on a protein destined for export from a cell is never located at the C-terminal end of the polypeptide chain. Why?

SOLUTION

This is because the signal sequence is required for passage of the polypeptide chain, *as it is being assembled*, through a membrane. The C-terminal portion is always synthesized last, and a signal sequence at this location could not allow the "threading" of the growing chain through the membrane.

INHIBITORS OF TRANSLATION

17.35. Why is puromycin able to function as an inhibitor of translation in both bacteria and eukaryotes?

SOLUTION

Puromycin is similar in structure to the 3′ end of aminoacyl-tRNAs (especially that of phenylalanyl-tRNA), with which it competes during translocation. The 3′ end of aminoacyl-tRNAs is the same in all organisms.

17.36. Chloramphenicol, when used as an antibiotic to treat bacterial infection in animals, can have side effects. What might contribute to these?

SOLUTION

Chloramphenicol blocks translation in bacteria by inhibiting peptidyltransferase of the large ribosomal subunit. It does not interfere with peptidyltransferase in the large subunit of eukaryotic ribosomes. However, the mitochondrion of animal cells contains ribosomes that are similar to bacterial ribosomes, and chloramphenicol can block protein synthesis in this organelle. This could contribute to the side effects of this drug when used in the treatment of animals.

CONTROL OF GENE EXPRESSION

17.37. Why is control of expression of the *lac* operon in *E. coli* said to be an example of *negative* control?

SOLUTION

This is because control of expression is brought about through modulating the effectiveness of a negatively acting agent, the repressor. In other words, expression of the *lac* operon is achieved through negating the effect of the repressor. In *positive* control systems, on the other hand, expression is achieved through the immediate effect of an agent to positively induce or increase expression.

17.38. What would be needed for the coordinate control of expression (at the transcriptional level) of a number of genes not located next to one another on the chromosome?

SOLUTION

Each gene would need to have common regulatory elements, promoters, operators, or both, associated with it. In this way, a single control factor, such as a repressor or positively acting substance, could influence all genes simultaneously.

17.39. Transcriptional control of globin synthesis in reticulocytes is not possible because transcription does not occur in these cells. Does this mean that the overall control of globin synthesis is completely lacking an aspect of transcriptional control?

SOLUTION

No. Prior to the formation of the reticulocyte from its precursor cells during the process of erythropoiesis (Chap. 1), there must have been a stage of preferential transcription of the globin genes to yield the globin-enriched mRNA of the reticulocyte.

Supplementary Problems

17.40. Decode the following RNA sequence into the corresponding amino acid sequence:

CAU AUU ACU CAU GAA CGU GAA

17.41. The following occurs at the start of the coding sequence of a eukaryotic mRNA:

GUG UUU UUU GUG UUU

For what amino acid sequence does it code?

17.42. The following segment of duplex DNA contains the region defining the start codon for a protein:

5′-GATGTCTCCT-3′

3′-CTACAGAGGA-5′

Identify the template strand.

17.43. Which is the most abundant RNA species in a cell?

17.44. DNA polymerase needs a primer on which to attach a new nucleotide unit for chain growth. Is a primer obligatory for RNA polymerase action?

17.45. For which is a promoter needed: initiation of RNA synthesis or of polypeptide synthesis?

17.46. The Pribnow box occurs about 35 bp upstream of the initiation codon. (*a*) True; (*b*) false.

17.47. The σ subunit of RNA polymerase is needed only for initiation of transcription. (*a*) True; (*b*) false.

17.48. Is the termination sequence for transcription near the 5′ or the 3′ end of the transcript?

17.49. Can monocistronic mRNA represent a primary transcription product of bacterial cells?

17.50. In which are enhancer sequences found, bacteria or eukaryotes?

17.51. Would phosphorylation of histones decrease or increase the net positive charge on such molecules?

17.52. Genes of eukaryotes are generally made up of introns and exons. Are these sequences both transcribed and translated?

17.53. Does splicing occur in DNA, RNA, or polypeptides?

17.54. When does capping of an RNA molecule occur, before or after splicing?

17.55. Can an exon contain an initiation codon?

17.56. Does rifampicin inhibit the transcription of histone genes?

17.57. In what species of RNA do codons and anticodons occur?

17.58. Why is the tRNA molecule so large, when the codon is only three nucleotide residues long?

17.59. tRNA is a single-stranded molecule, but much of its structure is double-helical. How is this possible?

17.60. Does the acceptor stem of a tRNA molecule contain the 3′ or the 5′ end?

17.61. Are there more species of aminoacyl-tRNA synthetase or tRNA in a cell?

17.62. Which of the subunits of a ribosome recognizes the nontranslated leader sequence of mRNA?

17.63. Subsequent to the formation of the 70S initiation complex in bacteria, at what steps of translation is energy consumed in the form of high-energy phosphate bonds?

17.64. Are termination codons associated with termination of transcription or termination of translation?

17.65. Is the methylated cap on eukaryotic mRNA attached to the RNA while it is in the nucleus or in the cytoplasm?

17.66. What is the first translational product of the insulin gene?

17.67. Does a signal peptidase cleave the nontranslated leader sequence of mRNA?

17.68. Does a bacterial repressor molecule bind to the promoter or operator site adjacent to a gene in DNA?

17.69. What is the difference between a repressor and a corepressor in relation to transcription?

Answers to Supplementary Problems

Chapter 1

1.15. The membranes of all living cells are *selectively* permeable to ions and other chemical species. This selectivity is in many cases linked to the supply of ATP (Chap. 10), and one aspect of cell death is low levels of ATP. In this state, the cells no longer exclude foreign compounds, such as toluidine dye.

1.16. Glutaraldehyde forms a Schiff base between side-chain amino groups of neighboring protein molecules, thus cross-linking them (Chaps. 3 and 4).

1.17. The arylsulfatase substrate *p*-nitrophenyl sulfate is used together with lead nitrate in a manner analogous to the Gomori reaction (Example 1.5).

1.18. Yes, problems would arise in interpreting the autoradiograph because the [^{3}H]glucose would not only be incorporated into glycogen but would also be metabolized via glycolysis (Chap. 11) to yield amino acids and fatty acids; these could appear in a whole array of cellular organelles.

1.19. Fragments of endoplasmic reticulum are transformed from lipid bilayer sheets, with attached ribosomes, into spherical vesicles. This is a result of the homogenization used in preparing the samples and also the tendency of lipid bilayers (Fig. 1-4) to *spontaneously* reseal.

1.20. It enables *separate control* over urea and pyrimidine synthesis (Chap. 15).

1.21. Incubate the reticulocytes with [^{3}H]leucine, which will be incorporated into proteins. Prepare electron microscope autoradiographs and count silver grains per cell and the number of polysomes. The latter appear as rosettes of five ribosomes in these cells. A statistical comparison between the number of polysomes and the amount of protein synthesized during the incubation time (proportional to the number of silver grains) indicates whether there are nonactive polysomes. In fact, many of the polysomes are inactive; i.e., they are "switched off" (see Chap. 17 for the control of protein synthesis).

1.22. (*a*) Mother. (*b*) If a defect exists in a mitochondrial gene, all progeny from that *female* will carry the defect. Several well-defined diseases resulting from such a defect have been described.

1.23. In fact, bacteria do not have mitochondria, but some types do have membranous intrusions into the cytoplasm called *mesosomes*. These are similar in function to the inner membrane of mitochondria (Chap. 14). The reason mitochondria are distinct from other membranous structures in higher cells is posssibly due to their evolutionary origin as intraeellular symbionts and to the fact that the spatial separation of functions leads to more advantageous (in terms of natural selection) control of the various metabolic processes.

1.24. (*a*) The release of peptidases, in particular, leads to tissue protein hydrolysis and hence breakdown. (*b*) Treatment is aimed at reducing inflammation with anti-inflammatory steroid drugs, which also serve to stabilize the lysosomal membranes.

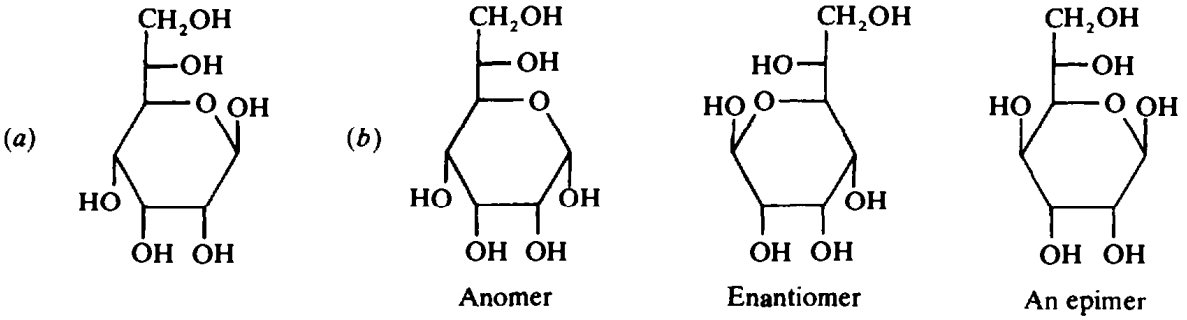
1.25. A fragment, usually the short arm, of chromosome 21 is translocated onto another chromosome; thus, there are three copies of a fragment of the short arm in any one cell. This is a relatively rare occurrence.

Chapter 2

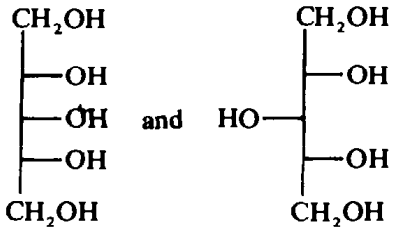
2.12. (*a*) Pentose, D-xylose, β; (*b*) pentose, D-ribose, β; (*c*) aminohexose, *N*-acetyl-D-glucosamine, β; (*d*) hexose, D-glucose, β.

2.13. (*a*), (*c*), (*d*), (*e*), (*f*), (*g*), (*h*), (*i*).

2.14.



2.15.

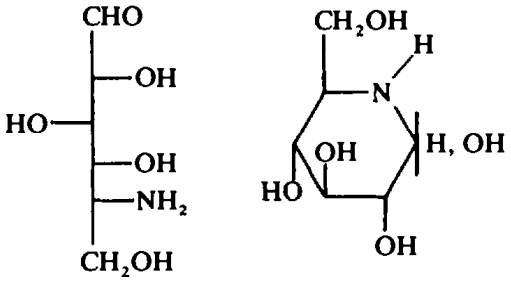


2.16. α (28 percent); β (72 percent).

2.17. 0.167 g cm^{-3}.

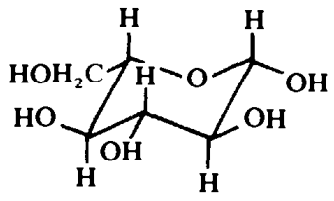
2.18. Fructose is (−).

2.19.

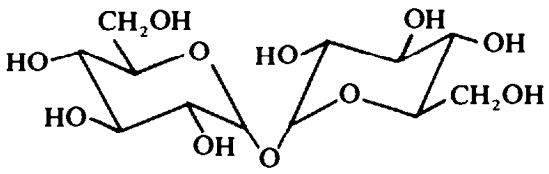


2.20. Six.

2.21. The substituents of β-L-glucopyranose are all on the opposite side of the Haworth structure from those of β-D-glucopyranose (i.e., they are in the positions occupied by the H's in β-D-glucopyranose). Therefore, from Example 2.13, it may be seen that in the 1C chair conformer, all the substituents of β-L-glucopyranose would be equatorial.

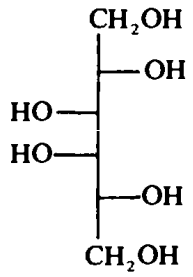


2.22.



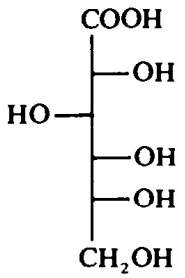
2.23. Erythritol is symmetrical, and both chiral carbon atoms are equivalent.

2.24.

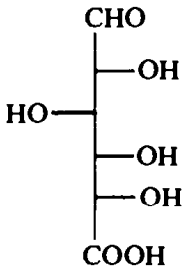


2.25.

(*a*)

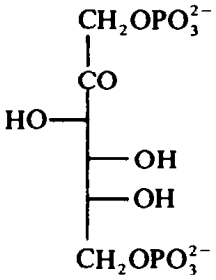


(*b*)

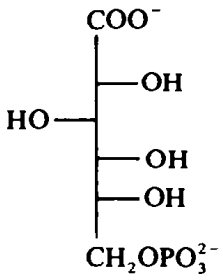


2.26. By inversion about C-5.

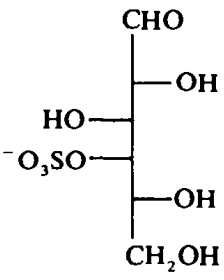
2.27.



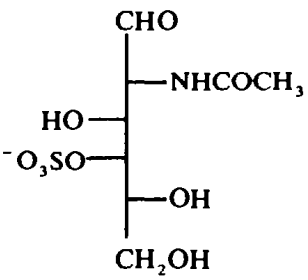
Fructose 1,6-diphosphate



6-Phosphogluconate



D-Galactose 4-sulfate



N-acetylgalactosamine 4-sulfate

2.28. No; L-fucose and L-rhamnose are deoxyhexoses.

2.29. Isomers.

Chapter 3

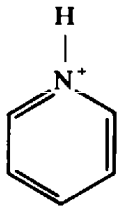
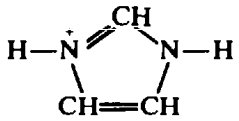
3.18.

(*a*) Serine, threonine (—OH); asparagine, glutamine ($—NH_2$); cysteine (—SH); tyrosine (—OH); aspartic acid, glutamic acid (—COOH); lysine, histidine, arginine (—NH).

(*b*) Serine, threonine (—OH); asparagine, glutamine (—C(=O)—); cysteine (—S—); tyrosine (—O—); aspartate, glutamate (—C(=O)—O$^-$); histidine, lysine, arginine (usually the unprotonated form)

3.19. Phenylalanine, tyrosine, tryptophan.

3.20. Glycine.

3.21. (*a*) $HO—CH_2—CH_2—SH$; (*b*)  (*c*) 

3.22. I.P. = 10.75.

3.23. I.P. = 7.6.

3.24. pH = 2.35.

3.25. (*a*) 1.61; (*b*) 1.0; (*c*) −0.15.

3.26. (*a*) 0.98; (*b*) 0.06; (*c*) −0.94.

Chapter 4

4.21. (*a*) Actin, myosin, tropomyosin, troponin, myoglobin; (*b*) collagen; (*c*) keratin.

4.22. 17,000.

4.23. (*a*) Anode.
(*b*) Cathode at pH 3.0; anode at pH 9.0.
(*c*) Cathode at pH 4.5; stationary at pH 9.5; anode at pH 11.0.
(*d*) Anode at pH 3.5, 7.0, 9.5.

4.24. Aldolase (first), serum albumin, hemoglobin, β-lactoglobulin, ribonuclease.

4.25. See Sec. 4.4 for these definitions.

4.26. (*a*) Surface: histidine, arginine, glutamine, glutamic acid; these are polar or charged.
(*b*) Interior: methionine, phenylalanine, valine; these are nonpolar. In addition, glutamine and uncharged histidine may be found in the interior if they can form hydrogen bonds.

4.27. (*a*) Nonpolar groups at the surface may participate in the binding sites for other molecules. (*b*) Charged groups in the interior may be important in the catalytic mechanisms of some enzymes.

4.28. A domain is an independently folded region of a protein; e.g., the NAD^+-binding domain of glyceraldehyde-3-phosphate dehydrogenase.

4.29. Such a conformation leads to clashes between atoms.

4.30. (*a*) Urea in high concentrations (6–8 mol L^{-1}) weakens hydrogen bonds and hydrophobic interactions.
(*b*) Urinary urea is very dilute (~0.2 mol L^{-1}).

4.31. (*a*) Hydrogen bonds, hydrophobic interactions, van der Waals interactions, charge-charge interactions.
(*b*) These act cooperatively.

4.32. (*a*) $p = 0.54$ nm, $d = 0.15$ nm.
(*b*) 153×0.15 nm = 22.95 nm.
(*c*) 153×0.35 nm = 53.55 nm.
(*d*) 153×0.36 nm = 55.08 nm.

4.33. (*a*) α helix.
(*b*) α helix (weak).
(*c*) disordered (charge repulsion).
(*d*) α helix (weak).
(*e*) α helix.
(*f*) probably disordered.

4.34. Hydrogen bonds, hydrophobic interactions, charge-charge interactions, van der Waals interactions.

4.35. Proline is too bulky to accommodate the close approach of the chains at every third residue.

4.36. Glycine is small enough to accommodate the close approach at every third residue.

4.37. (*a*) Insulin has two chains, formed by cleavage from a single-chain precursor; it is stabilized by disulfide bonds.
(*b*) Hemoglobin has four chains, two α and two β; it is held together by noncovalent interactions.
(*c*) Collagen has three chains and forms a triple helix; it is stabilized by hydrogen bonds and by additional chemical cross-links.

4.38. Maximal hydrogen-bond formation; all bond lengths and angles are normal; no clashes occur between atoms of the backbone (i.e., the α helix maps to a favorable region of the Ramachandran plot).

4.39. (*a*) Differences: H bonds in the α helix are formed between peptide groups of the same chain; in β structures, H bonds may be formed between different chains. H bonds are approximately parallel to the α-helix axis, and perpendicular to the direction of the chain in the β structures.
(*b*) Similarities: both the α helix and β structures are regular, repeating structures, are stabilized by hydrogen bonding, and map to favorable regions of the Ramachandran plot.

Chapter 5

5.19. (*a*) 10^{-3} mol L^{-1}; (*b*) 1 mol L^{-1}; (*c*) 10^{-2} mol L^{-1}.

5.20. (*a*) 24 percent; (*b*) 98.4 percent.
N.B.: The percentage of dimer increases as the equilibrium constant increases.

5.21. (*a*) 0.65.
(*b*) 0.025.
(*c*) 0.04.
(*d*) 0.74.
(*e*) 0.48.

5.22. The oxygen affinity of hemoglobin is lowered by the chloride ion strengthening the salt link. Thus, in the presence of chloride, more oxygen would be released.

$$Hb(O_2)_4 + Cl^- \rightleftharpoons Hb(Cl^-) + 4O_2$$

5.23. The oxygen affinity is reduced (p_{50} is 70 torr, and the arterial blood is only 70 percent saturated).

5.24. Polypeptides (*b*) and (*c*) would form triple helices, and (*b*) would be the more stable.

5.25. Since the polypeptide chains of collagen are almost completely extended, they cannot be stretched much farther without breaking covalent bonds. The α helix can almost double its length by breaking weak, noncovalent hydrogen bonds.

5.26. (*a*) 6.
(*b*) 4.

5.27. Hydroxyapatite binds in the gap between the collagen molecules in a way similar to the heavy-metal stains used in electron microscopy.

5.28 The proteases cleave the link and core proteins in the region nearest the hyaluronate, where there is least protection from the glycosaminoglycans.

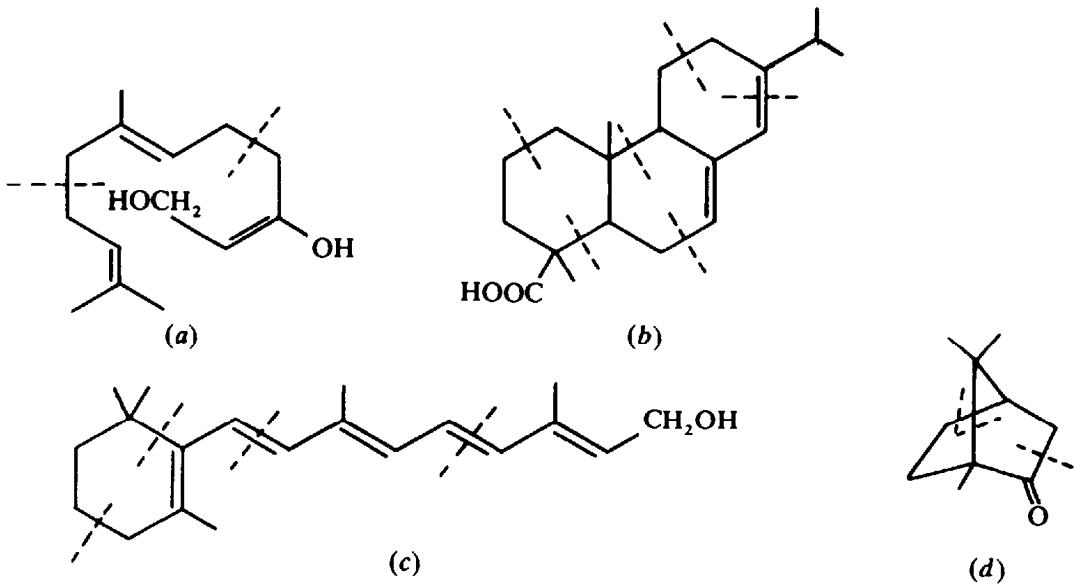
5.29. 9,400.

5.30. The intermediate filaments are cell-specific and do not change when the cell is transformed. Thus, by identifying the type of intermediate filament present in a tumor, the origin and characteristics of even a micrometastasis can be determined.

5.31. Dynein is probably carried out to the plus ends of axons in an *inactive* form, on vesicles transported along the microtubules by kinesin. Once at the axonal terminal, dynein is converted to an *active* form and can serve to transport vesicles back to the cell body.

Chapter 6

6.13.



6.14. (*a*) A lipid that contains carbohydrate; (*b*) 6-α-D-galactopyranosyl-β-D-galactopyranosyldiglyceride (a glycoglycerolipid) and galactosylceramide (a glycosphingolipid).

6.15. Because they are polymers of the two-carbon compound acetate.

6.16. (*a*) $CH_3(CH_2)_{12}COOH$.
(*b*) $CH_3(CH_2)_3CH{=}CH(CH_2)_7COOH$.
(*c*) $CH_3(CH_2)_5CH(OH)CH_2CH{=}CH(CH_2)_7COOH$.

6.17. 64 (24 if the three fatty acids are always different).

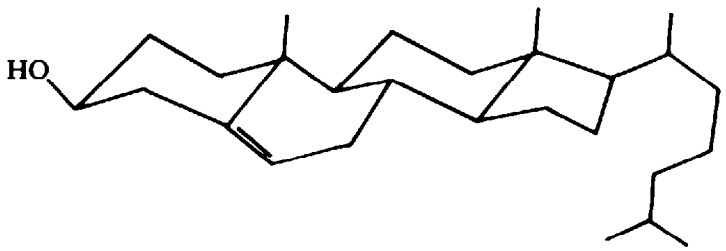
6.18.

(*a*) $CH_3(CH_2)_{14}CO\cdot OCH(CH_2O\cdot OC(CH_2)_{14}CH_3)(CH_2O\cdot OC(CH_2)_{14}CH_3)$ (Solid)

(*b*) $CH_3(CH_2)_{14}CO\cdot OCH(CH_2O\cdot OC(CH_2)_7CH{=}CH(CH_2)_7CH_3)(CH_2O\cdot OC(CH_2)_7CH{=}CH(CH_2)_7CH_3)$ (Liquid)

6.19. To the anode: PG, DPG, PS; stationary: PE.

6.20.



6.21. Phosphatidic acid.

6.22. 2.86×10^6.

6.23. Calculating the surface area of one red blood cell and multiplying this by the 4.74×10^9 cells yields a total area that is roughly only half the area found empirically. Therefore, the actual lipid content must be roughly twice that calculated for a monolayer, as would be the case for a surface bilayer.

6.24. PE has a smaller polar head than PC and packs in the concave inner surface better.

6.25. (*a*) Lowers; (*b*) raises; (*c*) raises transition temperature. (*a*) Increases; (*b*) reduces; (*c*) reduces mobility of phospholipids.

6.26. $1.04\ g\ cm^{-3}$.

6.27. 120.

6.28. 1-Propanol (fastest), 1,3 propanediol, propionamide, propionic acid, alanine.

6.29. The answer should be based on material provided by Example 6.9, Fig. 6-10, and the introduction to Example 6.12.

6.30. $3{,}300\ cpm\ min^{-1}$.

6.31. $35.5\ kJ\ mol^{-1}$.

6.32. (*a*) No effect; (*b*) transport inhibited.

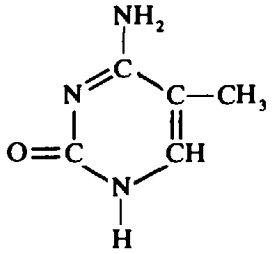
6.33. Ca^{2+} release is suppressed.

Chapter 7

7.35. Structure (*a*) is a pyrimidine, and (*c*) is a purine.

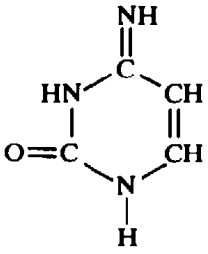
7.36. (*a*) RNA: adenine, guanine, uracil, and cytosine.
(*b*) DNA: adenine, guanine, thymine, and cytosine.

7.37.

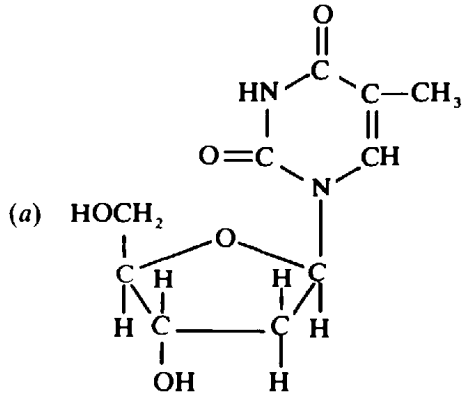


(*b*) Most DNAs contain very small amounts of 5-methylcytosine.

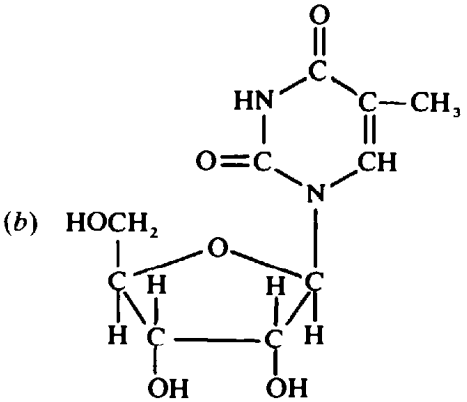
7.38.



7.39.



Deoxythymidine

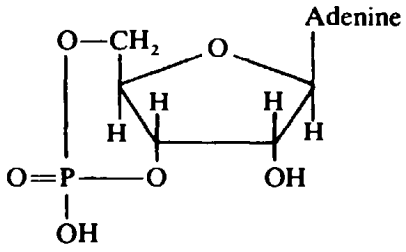


Ribothymidine

7.40. Adenosine.

7.41. A nucleotide contains a phosphate group; a nucleoside does not.

7.42.



7.43. (*b*) Deoxyguanosine monophosphate.

7.44. Because it contains uracil, which does not normally occur in DNA.

7.45. Less resistant.

7.46. *Sequence complementarity* refers to the matching of a sequence in a strand of DNA with that in another in terms of the ability of the two to form complementary base pairs.

7.47. The B form is a right-handed helix; the Z form is left-handed.

The repeating unit in the B form is a mononucleotide and a dinucleotide in the Z form.

The sugar phosphate backbone of the Z form follows a zigzag course, whereas that of the B form is relatively smooth.

7.48. At high ionic strength, positively charged counterions reduce the negative charge on the DNA, and this lowers the electrostatic repulsion within the structure, making it more stable.

7.49. Because the bacterial genome is smaller in size than the eukaryotic genome, nonrepeated sequences in DNA from the former will be present at a higher concentration. Renaturation is a bimolecular reaction and will proceed more quickly when the reactants are at a higher concentration.

7.50. Sequence complexity refers to the variety of sequences present in a genome.

7.51. The term *chromosome* refers to the structural unit within which the genetic material of a cell (usually DNA) is organized. The term *genome* describes a single complement of genetic material of the cell and could be made up of more than one chromosome.

7.52. Core histones are those around which the DNA in eukaryotic cells is wrapped to yield a nuclease-resistant particle. Linker histone is associated with the DNA that links these particles and is accessible to nucleases.

7.53. They contain a high proportion of positively charged amino acid residues that interact electrostatically with the negatively charged DNA.

7.54. (*a*) Positive supercoiling results from overwinding of the DNA helix; negative supercoiling results from underwinding of the helix. (*b*) Relaxation refers to the removal of the supercoiling, and (*c*) it can be achieved by nicking one of the two strands of DNA.

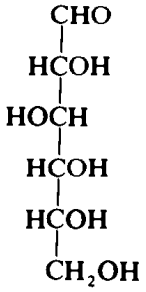
7.55. DNA contains deoxyribose as the sugar and thymine as one of the four bases. RNA contains ribose as the sugar and uracil as one of the four bases.

7.56. Because it is transcribed (formed) from a much larger portion of the genome than the other types of RNA.

7.57. (*c*) The sequence GTAATC would be the least likely to be cut because it does not have two-fold rotational symmetry.

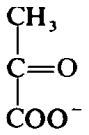
Chapter 8

8.14. (*a*) Glucose



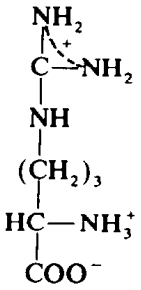
Glucose 6-phosphatase is a hydrolase, EC 3.1.3.9.

(*b*) Pyruvate



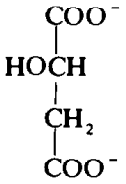
Lactate dehydrogenase is an oxidoreductase, EC 1.1.1.27.

(*c*) Arginine



Arginosuccinase is a lyase, EC 4.3.2.1.

(*d*) Malate



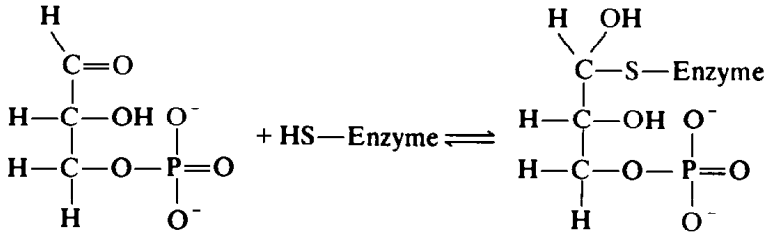
Fumarase is a hydratase, EC 4.2.1.2.

8.15. (*a*) Ethanol > methanol > propanol > butanol > cyclohexanol > phenol.

(*b*) Ethanol is the "best" substrate; methanol is relatively too small, while the rest are too large or too hydrophobic to be accommodated well in the active set.

8.16. (*a*) Thiohemiacetal.

(*b*)



8.17. (*a*) Mutases are members of the EC class 5, i.e., isomerases. The enzymes of class 5 catalyze geometrical structural changes *within* one molecule. When the isomerization consists of an intramolecular transfer of a *group*, the enzyme is called a mutase. Examples include phosphoglucomutase and chorismate mutase.

(*b*) These three enzyme types are all members of EC class 1: oxidoreductases. The enzymes of this class catalyze oxidation/reduction reactions, and the systematic name is *hydrogen donor:acceptor oxidoreductase*. The recommended name is *dehydrogenase* or, alternatively, *reductase*. *Oxidase* is used only in cases where O_2 is the acceptor. An example of a dehydrogenase is lactate dehydrogenase which catalyzes the formation of NADH and pyruvate from lactate and NAD^+.

An example of an oxidase is choline oxidase (EC 1.1.3.4), which catalyzes the formation of betaine and H_2O_2 from choline and oxygen.

8.18. In humans, the time spent by red blood cells flowing through capillaries of the lung alveoli is ~0.3 s. In that time, HCO_3^- in the plasma must reenter the red cells and be dehydrated to yield CO_2, which then diffuses across the membranes of the red cells and capillaries into the alveoli (and is expelled into the atmosphere). The spontaneous dehydration reaction is simply too slow.

8.19. A substrate analog on which an enzyme operates but which then covalently modifies the active site, permanently inhibiting it.

8.20. No. The ratio of the *free* substrate and product will be the same, but the substrate and product that are bound to the enzyme must be taken into account.

8.21. There is no single, simple answer to the question. Possible reasons are to: (1) provide the "correct" chemical environment for binding and catalysis, e.g., lower the pK_a of the group, (2) absorb energy of bombardment of the diffusive (thermal) motion of water and "funnel" it into the active site to enhance the catalytic rate, (3) allow for *control* of catalysis via conformational changes induced by effectors binding to other sites on the enzyme, (4) allow the fixing of enzymes in membranes or in large organized complexes, (5) prevent their loss by filtration through membranes, e.g., in the kidney.

8.22. DNA polymerase (Chap. 16), glycogen phosphorylase (Chap. 11).

Chapter 9

9.24.

$$v_0 = \frac{V_{max}[S]_0}{K_m + [S]_0}$$

$$v_0(K_m + [S]_0) = V_{max}[S]_0$$

$$\frac{v_0(K_m + [S]_0)}{[S]_0} = \frac{V_{max}[S]_0}{[S]_0}$$

$$\frac{v_0 K_m}{[S]_0} + v_0 = V_{max}$$

$$v_0 = V_{max} - \frac{v_0 K_m}{[S]_0}$$

9.25. A plot of the data pairs consisting of $(v_{0,i}/[S]_{0,i},\ v_{0,i})$ gives a straight line with a slope of $-K_m$ and ordinate intercept V_{max}.

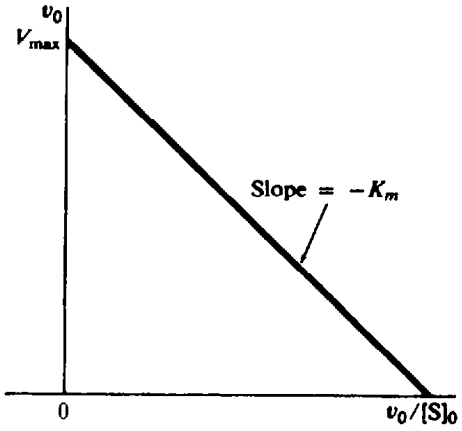


Fig. A-1 Eadie-Hofstee plot.

9.26. Multiply the Lineweaver-Burk equation by $[S]_0$ on both sides of the equals sign.

9.27. A plot of the data pairs $([S]_{0,i}, [S]_{0,i}/v_{0,i})$ gives a straight line with a slope of $1/V_{max}$ and abscissal intercept of $-K_m$.

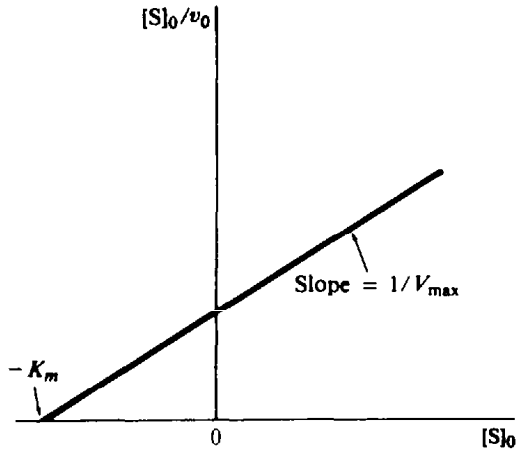


Fig. A-2 Hanes-Woolf plot.

9.28. (*a*) Remember to use reciprocals. (*b*) No.

9.29. (*a*) Separate the variables and integrate. Then

$$V_{max}t = [S]_0 - [S]_t + K_m \ln([S]_0/[S]_t)$$

(*b*) For a simple graphical procedure, the equation is rearranged to give

$$\frac{1}{t}\ln([S]_0/[S]_t) = -\frac{1}{K_m}\left(\frac{[S]_0 - [S]_t}{t}\right) + \frac{V_{max}}{K_m}$$

When the left-hand term is plotted versus $([S]_0 - [S]_t)/t$, the slope of the line is $-1/K_m$, the ordinate intercept is V_{max}/K_m, and the abscissal intercept is V_{max}.

9.30. (*a*) Separate the variables, noting that $[I] = [P] = [S]_0 - [S]$; then integrate directly:

$$V_{max}t = \left(1 + \frac{[I]}{K_I}\right)([S]_0 - [S]) + K_m \ln([S]_0/[S])$$

(*b*) Use the graphical analysis given in 9.29 above for a fixed $[S]_0$ value and a range of [I] values, to obtain a series of lines with the same slope but different abscissal and ordinate intercepts. A secondary plot of the reciprocal of the abscissal intercept versus [I] has the slope $1/K_I$.

9.31. Consider

$$\frac{d[S]}{dt} = -\frac{V_{max}[S]}{K_m\dfrac{[1 + ([S]_0 - [S])}{K_I} + [S]}$$

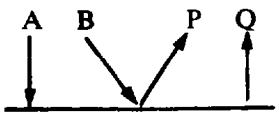
If $K_I = K_m$, then

$$\frac{d[S]}{dt} = -\frac{V_{max}[S]}{K_m\dfrac{[(K_m + [S]_0 - [S])]}{K_m} + [S]}$$

from which the result follows by direct integration, after cancellation of the K_m's in the numerator.

9.32. Use any of the graphical procedures. $K_m = 0.8\,\mu mol\,L^{-1}$; $V_{max} = 3\,\mu mol\,L^{-1}\,min^{-1}$; and from the expression $V_{max} = k_{cat}[E]_0$, $k_{cat} = 300\,s^{-1}$.

9.33. (*a*)



(*b*) Derive the rate expression using the King-Altman procedure. The relationships follow from arguments based on those used in Example 9.6.

9.34. This is an important question that has been addressed by many enzyme kineticists over the years. For the correct application of the Briggs-Haldane steady-state analysis, in a *closed* system, $[S]_0$ must be $\gg [E]_0$, where the $\gg$ sign implies a factor of at least 1,000. M. F. Chaplin in 1981 noted that the expression $v_0 = V_{max}[S]_0/(K_m + [S]_0 + [E]_0)$ yields, for example, only a 1 percent error in the estimate of v_0 for $[S]_0 = 10 \times [E]_0$ and $[S]_0 = 0.1\ K_m$; the expression thus applies under much less stringent conditions than does the simple Michaelis-Menten equation. In open systems $[S]_0$ can approximate $[E]_0$ and a steady state of enzyme-substrate complexes can pertain; computer simulation of both types of system is the best way to gain insight into the conditions necessary for a steady state of the complex.

9.35. (*a*) $\frac{1}{9}K_m$; (*b*) $\frac{1}{3}K_m$; (*c*) K_m; (*d*) $9\,K_m$

9.36. Differentiate Eq. (*9.36*).

9.37. (*a*) -7; (*b*) -10; (*c*) 81.

9.38. Yes. If $0 < n < 1$, the data conform to a negatively cooperative enzyme or binding protein; this is shown by using the analysis of Prob. 9.21, and in the present case $1/n > 1$, so $(81)^{1/n} > 81$.

9.39.

$$Y = \frac{K_1[X] + K_1K_2[X]^2}{1 + 2K_1[X] + K_1K_2[X]^2}$$

Chapter 10

10.12. Use the equations for activity and activity coefficient in Prob. 10.3. See Fig. A-3.

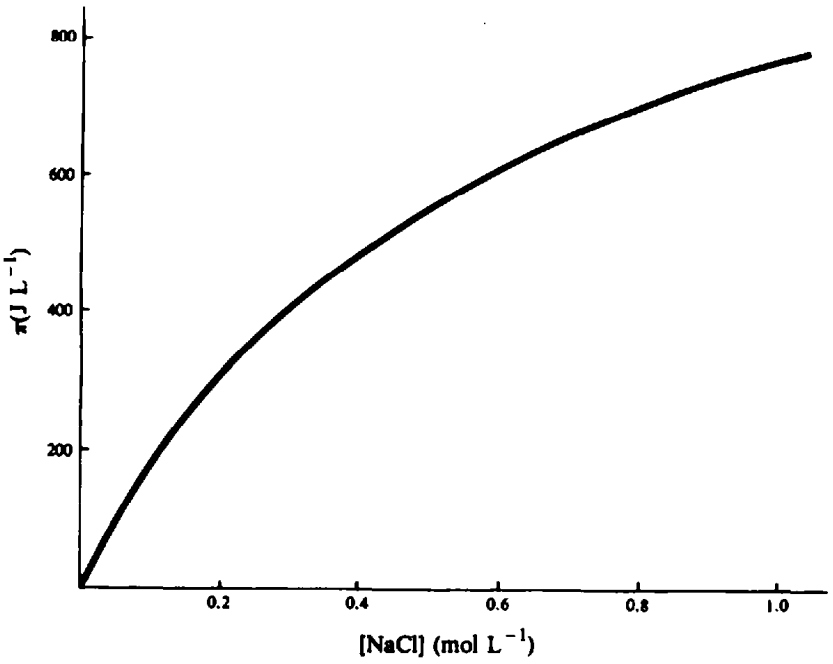


Fig. A-3

10.13. $\Delta G^{\circ\prime} = 219.9\ \text{kJ mol}^{-1}$.

10.14. Under biochemical standard-state conditions, the free energy of ATP synthesis is given by the expression

$$\Delta G^{\circ\prime} = -nF\Delta p$$

where n is the number of protons translocated per ATP molecule synthesized and Δp is the proton-motive force. Substituting values for $\Delta G^{\circ\prime}$ and n gives

$$\Delta p = -0.158\ \text{V}$$

Δp is given by the expression

$$\Delta p = \Delta\psi - \frac{2.3RT}{F}\Delta\text{pH}$$

which is Eq. (*10.35*) in which both sides have been divided by F.

Hence, $\Delta\mu = \Delta p \cdot \text{F}$

Therefore, $\Delta\mu = -15.25\ \text{kJ mol}^{-1}$

10.15. (*a*) 0.03; (*b*) 0.15; (*c*) 0.3; (*d*) 3.0.

10.16. Use Eq. (*10.10*).

10.17. (*a*) $\Delta G^{\circ\prime}_{\text{pH}=x} = \Delta G^{\circ} + RT\ln(10^{-x})$.

(*b*) Using this equation to calculate ΔG° for the value for $\Delta G^{\circ\prime}_{\text{pH}=7}$, we get $\Delta G^{\circ\prime}_{\text{pH}=6} = 23.2\ \text{kJ mol}^{-1}$.

10.18. (*a*) $\Delta S^{\circ} = 92.2\ \text{J K}^{-1}\ \text{mol}^{-1}$.

(*b*) The equilibrium constant at 323 K is 1.756×10^{-5}, if we assume that ΔH° and ΔS° are independent of temperature over the range 298–323 K.

10.19. (*a*) and (*b*).

10.20. At high $[\text{S}_i]$, $v_i = k_{\text{cat}_i}[\text{E}_i]$.

10.21. Consider a chain of enzymes E_1 to E_n,

$$\text{S}_1 \xrightarrow{\text{E}_1} \text{S}_2 \xrightarrow{\text{E}_2} \text{S}_3 \cdots\cdots \text{S}_n \xrightarrow{\text{E}_n} \text{S}_{n+1}$$

with various values of $V_{\max}$ and K_m, etc. The system is assumed to be in a steady state with constant flux of v_g through it. If, in a hypothetical experiment, we increased all enzyme concentrations (an enzyme's activity is assumed to be proportional to its concentration) by the same small factor α (i.e., $\Delta[\text{E}_i]/[\text{E}_i] = \alpha$) the fractional change in v_g would be exactly α:

$$\frac{\Delta v_g}{v_g} = \alpha$$

Now, provided α is very small, $\Delta v_g/v_g$ can be the sum of all the *individual changes* that would be caused by alterations to each of the *separate* enzymes. For the ith enzyme, the control coefficient is defined by C_i (see Prob. 10.20); therefore,

$$C_i = \left(\frac{\Delta v_g}{v_g}\right)_i \Big/ \frac{\Delta[E_i]}{[E_i]}$$

$$= \left(\frac{\Delta v_g}{v_g}\right)_i \Big/ \alpha$$

$$\alpha C_i = \left(\frac{\Delta v_g}{v_g}\right)_i$$

The summation of this equation gives

$$\frac{\Delta v_g}{v_g} = \sum_{i=1}^{n} \left(\frac{\Delta v_g}{v_g}\right)_i = \alpha \sum_{i=1}^{n} C_i$$

But, from the first equation, above, $\Delta v_g / v_g = \alpha$; therefore $\sum_{i=1}^{n} C_i = 1$.

Chapter 11

11.12. During glycolysis, glyceraldehyde 3-phosphate is converted to 1,3-bisphosphoglycerate and the equilibrium of the adenylate kinase reaction lies in favor of 3-phosphoglycerate, so the metabolites are "drawn" through the pathway of reactions.

11.13. 2,3-Bisphosphoglycerate is synthesized from 1,3-bisphosphoglycerate, a reaction catalyzed by 2,3-bisphosphoglycerate synthase. The enzyme 2,3-bisphosphoglycerate phosphatase catalyzes the hydrolysis of 2,3-bisphosphoglycerate to 3-phosphoglycerate.

11.14. Two of the carbon atoms are converted to carbon dioxide, and the other four carbon atoms yield two molecules of ethanol.

11.15. Four.

11.16. It can be incorporated into glycogen, used in the pentose phosphate pathway, or hydrolyzed to glucose.

11.17. The gut microbiota of termites synthesize and secrete the enzyme cellulase, which hydrolyzes the cellulose.

11.18. (*a*) Alcoholic fermentation operates: the glucose is converted to pyruvate, which forms CO_2 and ethanol, and NAD^+ is regenerated. (*b*) The inhibitor would block glycolysis, and the yeast would die.

11.19. (*a*) Insulin promotes the uptake of glucose from the blood to all cells. Excess glucose is converted to glucose 6-phosphate (via glucokinase) in the liver and is then stored as glycogen.

(*b*) Glucagon mobilizes the breakdown of glycogen in the liver, which produces glucose to maintain an adequate concentration in the blood.

11.20. Blood glucose concentrations would remain high after meals, and much of the glucose would be excreted in the urine. Glucose metabolism would operate, therefore, at a lower level than normal.

11.21. The pentose phosphate pathway would produce the necessary NADPH, and the ribose 5-phosphate would be converted into the glycolytic intermediates fructose 6-phosphate and glyceraldehyde 3-phosphate.

11.22. Glycolysis, the pentose phosphate pathway, the production of glycogen, and the formation of glucose (glucose-6-phosphatase).

Chapter 12

12.12. (*a*) Citric acid cycle: two; (*b*) glyoxylate cycle: zero.

12.13. Fumarate is the substrate for fumarase, whereas maleate is not. The cell could not use maleate as a carbon source.

12.14. None. Red blood cells do not have mitochondria and so do not have the enzymes to operate the citric acid cycle.

12.15. The inhibitor would prevent the citric acid cycle from operating, but in germinating plant cells the glyoxylate cycle would be unaffected. Thus, energy production would be decreased in both cells, but their ability to synthesize glucose would be unimpaired.

Chapter 13

13.17. Cholesterol is the precursor of bile salts, and their secretion into the intestine is stimulated after eating food and during the digestion and absorption of triglyceride.

13.18. Methyl groups are utilized in the de novo synthesis of phosphatidylcholine, a process that decreases the availability of 1,2-diacylglycerol for triacylglycerol synthesis.

13.19. The complete oxidation of 1 mole of palmitic acid to CO_2 and H_2O produces NADH and $FADH_2$ by β-oxidation and by citric acid cycle activity. ATP synthesis is coupled to the oxidation of NADH and $FADH_2$ produced in these processes (Chap. 14). Per mole of palmitoyl-CoA oxidized, 131 moles of ATP are synthesized. But, two high energy bonds of ATP are used in the formation of palmitoyl-CoA, so the net ATP production is 129.

13.20. (*a*) Acetoacetate is completely oxidized by entry of two molecules of acetyl-CoA into the citric acid cycle; 22 moles of ATP per mole of acetoacetate are synthesized from NADH and $FADH_2$ produced by the cycle. Two additional moles of ATP arise from two moles of GTP, also produced in the citric acid cycle. So, the total is 24.

(*b*) 3-Hydroxybutyrate is oxidized by β-oxidation to acetyl-CoA, which then enters the citric acid cycle; 25 moles of ATP per mole of 3-hydroxybutyrate are synthesized from the NADH and $FADH_2$ produced, and 2 moles of ATP from GTP are also synthesized. So, the total is 27.

13.21. (*a*) Palmitic acid and 3-linoleoylglycerol; (*b*) arachidonic acid.

13.22. Glucose is oxidized in the pentose phosphate pathway (Chap. 11), to produce NADPH that can enter cholesterol synthesis. One molecule of glucose is required per molecule of lanosterol synthesized. (The reactions that convert lanosterol to cholesterol are outside the scope of Chap. 13.)

13.23. Via β-oxidation.

13.24. Glucose is converted to dihydroxyacetone phosphate in glycolysis (Chap. 10), which is then reduced to glycerol 3-phosphate by NADH-dependent glycerol-3-phosphate dehydrogenase.

13.25. Acetyl-CoA carboxylase exists in the cytosol in two forms, one a monomer and the other a polymer. Citrate converts the inactive form (monomer) to the active form (polymer), and long-chain acyl-CoA is a feedback inhibitor that converts the active form back to the inactive form.

13.26. Transamination reactions produce pyruvate, and deamination followed by degradation of the carbon skeleton of ketogenic amino acids produces acetyl-CoA (see Chap. 15).

13.27. Acyl-CoA dehydrogenase, enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, thiolase, and Δ^3-*cis*-Δ^2-*trans*-enoyl-CoA isomerase.

13.28. Fat depots provide fatty acids for cellular fuel. The oxidation of fatty acids produces NADH and $FADH_2$, which are oxidized by oxygen (via the respiratory assemblies in the inner membrane of mitochondria, Chap. 14) with the concomitant production of water.

13.29. Fatty acid oxidation would be inhibited because acyl groups could not be transported into the mitochondrial matrix. This would restrict long-term muscle activity.

13.30. Yes, because platelet aggregation depends on the release of thromboxanes from the endothelial cells of blood vessels. Aspirin inhibits cyclooxygenase and thus synthesis of prostaglandins and thromboxanes is inhibited.

13.31. Triacylglycerol would be synthesized from dietary carbohydrate and protein.

13.32. *Cytosolic* 3-hydroxy-3-methylglutaryl-CoA synthase produces the 3-hydroxy-3-methylglutaryl-CoA for cholesterol synthesis, and *mitochondrial* HMG-CoA synthase produces the HMG-CoA, which is converted to acetoacetate.

13.33. Cholesterylesters arise from the activity of acyl-CoA:cholesterol acyltransferase, which catalyzes the formation of the esters from acyl-CoA, and also from the activity of lecithin:cholesterol acyltransferase, which catalyzes the formation of the ester from phosphatidylcholine.

13.34. Depot fat is used, and fatty acid β-oxidation in the liver increases. This provides the acetyl-CoA for the synthesis of ketone bodies.

13.35. Long-chain fatty acids are relatively insoluble in polar media, but they still form salts, some of which act as detergents (soaps) and thus associate with hydrophobic lipids in the lipoproteins.

13.36. Cells that synthesize prostaglandins, thromboxanes, and leukotrienes.

Chapter 14

14.11. (*a*) 15; (*b*) 3; (*c*) 36; (*d*) 16.

14.12. The P/O ratio would decrease.

14.13. Approximately 14.

14.14. Review Prob. 14.8, and think about the metablic sources of NADH and $FADH_2$. Uncoupling agents, such as 2,4-dinitrophenol, render the inner mitochondrial membrane permeable to the protons extruded during electron transport. The energy released as these protons return to the mitochondrial matrix leads to an increase in body temperature; this is relieved by sweating. Because, under these conditions, ATP synthesis is no longer completely coupled to electron transport, the oxidation of NADH and $FADH_2$ must increase to maintain a supply of ATP to the cell. As these compounds are produced during catabolism of carbohydrates, lipids, and proteins, there is an increased turnover of these body constituents, leading to a loss of body weight.

14.15. On the basis of their visible absorbance spectra. Many textbooks give examples of such spectra.

14.16. Apply the crossover theorem, with reference to the figure in Example 14.5. The proportions of the iron-sulfur proteins in complexes I, II, and III with iron atoms in the Fe^{3+} and Fe^{2+} states can, in principle, be estimated from the electron spin resonance signals of the preparation of mitochondria. In the case of mitochondria oxidizing NADH in the presence of rotenone, application of the crossover theorem reveals that the iron atoms in the iron-sulfur proteins in complex I will become more reduced, while those in complex III will become more oxidized, in comparison to their states in the absence of rotenone. In the case of KCN-treated mitochondria, both sets of iron-sulfur proteins will become more

reduced. With antimycin A-treated mitochondria, the complex I iron-sulfur proteins will become more reduced, while those in complex II may, or may not, depending on the site of action of antimycin A within this complex.

14.17. 114 kJ mol^{-1}.

14.18. Nitrite converts some ferrohemoglobin to ferrihemoglobin. The latter is able to bind cyanide, preventing it from inactivating cytochromes *a* and a_3 by binding to their ferriheme rings.

14.19. (*a*) Oxidative phosphorylation is more efficient than substrate-level phosphorylation.

(*b*) The NADH needed for lactate production will be oxidized to NAD^+ in oxidative phosphorylation.

Chapter 15

15.18. A person on this diet would be in negative nitrogen balance and could not sustain normal metabolism.

15.19. Yes, because cheese contains animal protein.

15.20. (*a*) Chymotrypsin and elastase; (*b*) Elastase, trypsin, and chymotrypsin; (*c*) Pepsin, chymotrypsin, and trypsin.

15.21. Glutaminase, aspartate aminotransferase, TCA cycle enzymes, citrate lyase, malate dehydrogenase, and $NADP^+$-linked malate enzyme.

15.22. The concentration of free arginine in the liver, and hence its release into the blood, is insufficient to sustain growth in children, so a dietary source of arginine is essential.

Chapter 16

16.34. (*a*) In unidirectional replication, a single replication fork generated at the origin moves in a unique direction along the DNA. (*b*) The rolling-circle mechanism is an example of undirectional replication.

16.35. In embryonic cells, replication is initiated at sites (origins) along the DNA that are more closely spaced than in terminally differentiated cells. Thus, the distance over which each replication fork travels to replicate the complement of DNA is very short, and the overall process is rapid.

16.36. The $5' \rightarrow 3'$ exonuclease and polymerase activities are utilized on the lagging strand arm to remove the RNA primers and replace them with DNA, respectively. The $3' \rightarrow 5'$ exonuclease activity is used in proofreading during this process.

16.37. At the origin of replication.

16.38. It could aid unwinding in one of two ways; by introduction of negative supercoiling, or removal of positive supercoiling ahead of the replication fork.

16.39. No.

16.40. There must be present a 3′ hydroxyl and a 5′ phosphate, one at each end of the chains to be joined.

16.41. This refers to the $3' \rightarrow 5'$ exonuclease activity associated with the enzyme. It acts to remove mismatched nucleotides incorporated infrequently by the polymerase.

16.42.

Origin of replication

↓ DNA gyrase + SSB protein

Stabilized, unwound
DNA at origin

↓ Primosome + primase + rNTPs +
DNA polymerase III
+ dNTPs

Leading strands extending
from primer and away from
origin on opposite templates

↓ Movement of
primosome away
from origin on
opposite strands

Priming and synthesis of
lagging strands. Generation
of replication bubble.

16.43. A *dna* mutation is one that alters a gene product directly involved in the replication process.

16.44. Yes, to the extent that replication is dependent on the availability of dTTP.

16.45. It would block initiation of replication at the origin.

16.46. No.

16.47. DNA polymerase I (with its associated $5' \rightarrow 3'$ exonuclease activity) and DNA ligase.

16.48. No.

16.49. It must have an origin of replication that is functional in the host cell.

16.50. The primers. These define the ends of the sequence that is to be amplified.

Chapter 17

17.40. His-Ile-Thr-His-Glu-Arg-Glu.
(By using the single letter notation for amino acids (Table 3.1), this sequence reads: HI THERE!)

17.41. Met-Phe-Phe-Val-Phe.

17.42. The lower strand.

17.43. Ribosomal RNA (rRNA).

17.44. No. The RNA polymerases can start chains de novo.

17.45. RNA synthesis.

17.46. (*b*).

17.47. (*a*).

17.48. Near the 3′ end.

17.49. Yes. (The *lac* repressor gene of *E. coli* would give rise to monocistronic mRNA.)

17.50. Eukaryotes.

17.51. Phosphorylation would *decrease* the net positive charge.

17.52. No. Introns and exons are both transcribed, but only exons are translated.

17.53. RNA.

17.54. *Before* splicing.

17.55. Yes. The first exon (closest to the 5′ end) of mRNA must contain the initiation codon.

17.56. No. Histones occur in eukaryotes, and rifampicin blocks transcription only in bacteria.

17.57. Codons occur in mRNA; anticodons, in tRNA.

17.58. The tRNA molecule must not only recognize the codon, but also recognize specific regions of the ribosome, elongation factors, and the appropriate aminoacyl-tRNA synthetase, each of which involves cooperative interactions through noncovalent bonds.

17.59. The single tRNA chain folds back on itself to allow the formation of several segments of double helix through intrastrand base pairing.

17.60. It contains both.

17.61. There are more species of tRNA.

17.62. The small ribosomal subunit.

17.63. At two steps: the binding of each charged tRNA and translocation after the formation of each peptide.

17.64. Termination of *translation*.

17.65. Capping of the RNA occurs *before* splicing and must occur in the nucleus.

17.66. Preproinsulin.

17.67. No.

17.68. It binds to the operator site.

17.69. A *repressor* molecule can bind to an operator and block transcription; a *corepressor* is a compound which binds to a protein to give a functional repressor, which can then bind to an operator.

Index

The letters *d*, *s*, and *t* following page numbers stand for diagram, structure, and table, respectively.

$[\alpha]_D^T$, 26, 34, 52
Abietic acid, 185*s*
Absorption, of dietary lipid, 391
Absorptive cells, 17
ACAT, acyl-CoA: cholesterol acyltransferase, 366*d*, 390
Acetal, 32*s*, 43*s*
Acetaldehyde, 232*s*, 322*s*
Acetate, as precursor of lipids, 154*s*
Acetic acid:
  oxidation energetics, 304
  as pH buffer, 72*s*
  pKa, 59
Acetoacetate, 370*s*, 371, 373*sd*, 396
  utilization, 374*d*
Acetoacetyl-CoA, 374, 388*s*
Acetogenins, acetate as precursor, 154*s*
Acetone, 370*s*, 373*sd*
Acetyl phosphate, energetics of, 309*d*
Acetyl transacylase, 377*t*
Acetyl-CoA, 370, 374, 374*d*, 388*s*, 395, 396
  acetyltransferase, 373*t*
  in the citric acid cycle, 345*s*, 346*s*, 350*s*
  in the glyoxylase cycle, 356*s*
  sources in metabolism, 351
Acetyl-CoA carboxylase, 375, 394
Acetylcysteinyl-polypeptide, 376*d*
N-Acetyl-D-galactosamine, 42*s*, 47
N-acetylgalactosaminidyl group, 383
N-Acetyl-D-glucosamine, 42*s*, 47
N-Acetyl-β-D-glucosamine, 47, 126, 245
N-Acetyl-D-glucosamine transferase, 149
N-Acetylmuramic acid, 245
N-Acetylneuraminic acid, 42*s*
N-Acetylneuraminidyl group, 383
Acid:
  acid-base behavior, 58, 59
  Brønsted definition of, 58
  conjugate acid, 58, 59
  weak acids, 58, 59
Acid anhydride bond, 231*t*
Acid catalysis, 233
Acid deoxyribonuclease, 11*t*
Acid fuchsin, 19
Acid hydrolase, 9, 10*d*
Acid hydrolases, in lysosomes, 20*d*
Acid phosphatase, 3*d*, 11*t*
Acid phosphodiesterase, 11*t*
Acid ribonuclease, 11*t*
Acid-base catalysis, general, 232, 233
Acidophilic, stain, 19
Aconitase, 13*t*, 346, 347*d*
  iron-sulfur protein, 347
  location, 13*t*
cis-Aconitase, 347*s*
Acridine, 473, 474*s*
Acrosomal process, 108
Actin, 113*d*, 132, 133, 139, 150
  ADP with, 132
Actin (*Cont.*):
  assembly, 132*d*, 134
  ATP with, 132
  barbed end, 133, 134*d*
  capping proteins and, 135
  critical concentration, 132, 133, 134*d*
  double helix of, 113, 131, 132
  (minus) end, 133*d*
  (plus) end, 133*d*
  F-, 134*d*
  G-, 134*d*
  lag period in polymerization, 132
  nucleus, 132*d*
  pointed end, 133, 134*d*
  polarity, 133
  polymer, 131
  polymerization, 132*d*
  thymosin binding, 134
  treadmill, 133
  various forms, 136*d*
Actin double helix, 113
Actin filament, 12
Actin-binding protein, 134, 135, 140, 150
α-Actinin, in Z-disc, 135, 137
Actinomycin D, 473, 499, 500, 513
Activation energy, 234, 237, 247
Activation of amino acids, for protein synthesis, 501*d*
Active site, 238
  of enzyme, 233, 234
Active transport:
  examples, 179*d*
  group translocation, 179*d*
  across membranes, 177, 178*t*
  primary active, 179*d*
  secondary active, 179*d*
Activity coefficient, 251
  definition, 305*d*
Actomyosin, 17, 144
Actomyosin, conformational change, 139*d*
Actomyosin, power stroke, 138
Acyl transacylase, 377*t*
Acyl-CoA: cholesterol acyltransfer (ACAT), 366*d*, 390
Acyl-CoA dehydrogenase, 370, 372*t*
Acyl-CoA derivatives, 370
Acyl-CoA desaturase, 397
Acyl-CoA synthetase, 369, 378
Acyl-CoA transferase, 378
Acyl-enzyme intermediate, 241*d*, 242*d*
Acylate, fatty acid salt, 155
Acylimidazole, 242
Adair equation, 268, 269, 271, 289
Adair, G. S., 268
Adamantane, 245*s*
Adenine, 200*s*, 205*s*, 458
Adenosine, 201*s*, 202
Adenosine deaminase, 456
Adenosine 5′-phosphate, AMP, 202*s*, 203
Adenosine triphosphate, ATP, 203
S-Adenosylmethioneine (SAM), 450*d*
Adenyl cyclase, in glycogenolysis, 337*d*
Adenylate cyclase, 392
Adenylate kinase, 436
  location, 13*t*
  thermodynamics of, 300*d*
Adenylic acid, AMP, 203
Adiabatic calorimeter, 291*d*, 304
Adiabatic (closed) system, definition, 290, 291
Adipocyte, 16*d*, 17
Adipose tissue, 392
ADP:
  in glycolysis, 311
  in muscle contraction, 137, 138, 139*d*
ADP-3′-phosphate, in CoA, 345*s*, 346
ADP-ribosylation, 508
Adrenal cortex, 336
Aglycone, 43
Alanine, 17, 67*s*, 238, 423*t*
  aminotransferase, 422*d*
  structure of, 54*t*
β-Alanine, 56, 70*s*
D-Alanine, 240*s*
L-Alanine, 240*s*
Alanine racemase, 241
Alanine transaminase, 241
  aminotransferase, 241
  pyridoxal phosphate in, 241
Alcaligenes faecalis, 227
Alcohol dehydrogenase, 249, 322*d*
  horse-liver, 289
Aldehyde, 3
  groups, 25, 42, 43
Alditol, 40
Aldohexose, 25, 28
Aldol, 125*d*
Aldolase:
  in glycolysis, 314*d*
  molecular weight, 83*t*
Aldonate, 41
Aldonic acid, 41
Aldopentose, 26, 28
Aldose, 25
Algae, blue-green, 419
Allolactose, 508
Allose, 28*s*
Allosteric:
  coefficient, 273*d*
  effect, 270
  effector site, 261
  interactions in hemoglobin, 118
Allysine, 124, 125*d*, 126
Altman, C., 258
Altmann, R., 12
Altrose, 28*s*
Amanita phalloides, 494
α-Amanitin, 494, 500
Amidino, 230*s*

Amine oxidase, 434
Amino acid, 17, 23, 53*s*–75
  acid-base behavior, 56–65
  activation, 501*d*
  analysis, 65
  essential, 419
  fates of carbon skeleton, 431*d*
  formal titration of, 62
  isoelectric point, 62, 64, 65
  lifetime in plasma, 431
  ninhydrin reaction with, 65*s*
  pKa values, 63*t*
  plasma concentration, 431
  with prototropic side chains, 63
  reaction with ninhydrin, 65
  site of synthesis, 17
  structure of, 53*s*
  synthesis, 419
  transport, 430
  tritiated, 4
  as zwitterion, 61–65
D-Amino acid oxidase, 12
L-Amino acid oxidase, 12
Amino sugars, 42
2-Amino-2-deoxy-D-glucose, D-glucosamine, 42
Aminoacyl-tRNA, 504*d*, 507*s*
Aminoacyl-tRNA synthetase, 501*d*, 514, 518
γ-Aminobutyrate (GABA), 70*s*, 455
δ-Aminolevulinate, δ-aminolevulinic acid, (ALA), 451*s*
δ-Aminolevulinate synthase, ALA synthase, 451, 452
  eIF-2 in regulation, 452
  inhibition by hemin, 452
  protein kinase and, 452
  regulation, 452
Aminopeptidase, 248, 430, 505
6-Aminopurine, 200
Aminotransferases, 421, 422, 423*d*
Ammonia, 419, 420
  equilibrium with ammonium, 434*d*, 435
AMP, 240*d*
  adenosine 5′-phosphate, 202*s*, 203*s*
  binding to phosphorylase a, 115
3′,5′-AMP, cAMP, 227
Amphibolic reactions, citric acid cycle, 354, 355*t*
Amphipathic, 165
Amphiphilic lipids, 165, 167
Amplification,
  of DNA, 464*d*
  of metabolic signals, 301
  in PCR, 480
Amylo-(1,4 → 1,6)-transglycosylase, branching enzyme, 328*d*
Amylopectin, 46
  structure, 330
Amylose, 46, 51
  structure, 330
Anaphase, 142, 143, 144
Anaplerotic reaction:
  C1 compounds, 449
  of the citric acid cycle, 355*t*
Andrews, P. R., 244, 245
Androgen, 393*d*
Anemia, 18, 131
Anhydride(s), 242, 243
Animal, 1
Anion channel, 136, 137*d*
Anion transport, in erythrocytes, 16
Ankyrin, 136, 137*d*
Annealing, 213
  DNA, 477*d*, 479, 480*d*
Anticodon, 500, 501*d*
Anticompetitive inhibition, 208*d*, 254, 255*t*, 264
Antilipolytic hormone, insulin, 394
Antimicrobial agents, 245
Antimycin A, inhibition of electron transport chain, 406*d*
Antisense strand, template, 489
Aphidicolin, 474, 475*s*
Apo proteins of lipoproteins, AI, AII, B, CI, CII, CIII, D, 169
Apoferritin, 453
Apolipoprotein B, 366
Apolipoprotein C, 365
Apolipoprotein E, 366
Apoprotein, 81
Arabino-hexulose, D- (fructose), 30, 31
Arabinose, 28*s*, 47
Arachidonic acid, 156*s*, 385*sd*, 386*d*, 387*d*, 398
  arachidonate, 376, 377, 384, 385*s*
  synthesis of, 377*d*
Arginine, 230*s*, 238, 423*t*, 424, 430*s*, 436*s*, 436
  amidinohydrolase, 230
  structure of, 56*t*
Argininosuccinate, 436*s*
Argininosuccinate lyase, argininosuccinase, 249, 424, 436*d*
Argininosuccinate synthetase, 424, 436
Arginyl, staining, 2
Aromatic compound, 2
Arrhenius, S., activation energy, 237, 238
Arthritis, 129
Arylhydrazinopyrimidine, 474
Arylsulfatase, 11*t*, 23
  A, 383*d*
Ascorbate, 505
  and prolylhydroxylase, 123*d*
  as secondary redox mediator, 406
Asparagine 70*s*, 423*t*
  structure of, 54*t*
Aspartate, 70*s*, 283, 422*s*, 423*t*
  aminotransferase, 422
  structure of, 55*t*
Aspartate carbamoyltransferase, 229, 266
  active site, 266
  allosteric site, 266
  aspartate as substrate, 229
  ATP and, 266
  ATP binding, 115
  carbamoyl aspartate in, 229
  carbamoyl phosphate as substrate, 229
  catalytic subunit, 115
  CTP and, 266
  CTP binding, 115
  effector molecules, 229, 266
  Escherichia coli (*E. coli*) form, 266
  inhibitor(s), 229
  subunit structure, 115, 116*d*
  subunits of, 229
  X-ray diffraction of, 115
Aspartate semialdehyde, 283
Aspartic acid:
  ionization of, 71*s*
  isoelectric point of, 71
Aspartokinase, 283
Aspirin, 398*s*
Association binding constant, 268
Astral microtubule, 143*d*
Atmospheric nitrogen, 419
ATP, 9, 21, 240*d*
  adenosine triphosphate, 203
  bioenergetic properties, 298
  free energy hydrolysis, 402
  in glycolysis, 311
  and hexokinase, 237
  in muscle contraction, 137–139*d*
  production in glycolysis, 314
  in pyruvate carboxylase reaction, 353*d*
  structure, 203*s*
ATP and actomyosin, 139
ATP synthase, 13*t*, 402
  in bacteria, 411
  in chloroplasts, 411
  energetics of, 309
  location, 13*t*
  mechanism, 413*d*, 414*d*
  in mitochondria, 411
  subunit structure, 412
ATP translocase, 241
  inhibitors, 415
  in mitochondria, 415
ATP utilization in macrophages, 21
ATPase, 237*s*, 289
  in muscle contraction, 137–139*d*
Atractyloside, inhibitor of ATP translocase, 415
Attachment glycoproteins, 130
Attenuator sequence, mRNA, 512
Autonomously replicating sequences (ars), 472
Autophagic vacuole, 19, 20*d*
Autophagosome, 20*d*
Autophagy, 9, 20*d*
Autoradiography, 4, 23
Autosomes, 15
Average life span of a cell, 21
Avogadro's number, 22
Axial configuration, 36*s*
Axon, 140, 141*d*
Axoneme, in cilia, 141, 142*d*
Azide, inhibition of electron transport chain, 406*d*
Azure I, 1

Bacteria, 12, 21, 419
  bacterium, 1
  phagocytosis of, 21
Bacterial cell wall, 11*t*
  polysaccharides, 245
Bacterial plasmid, 477
Bacteriophage, genome size, 215, 216*t*
Bacteriophage T4, virus, 110*d*, 111*d*
Bacteriophage T5, tobacco mosaic, 110
Bacteriophage T6, polio, 110
Barbed end:
  α-actinin binding to, 137
  of actin, 136
Basal body, 142
Base, definition of, 59

Base catalyst, 233
Base pairs in genes, 149
Base-pairing geometry, 210
Basement membrane, 122*t*
  binding, 131*t*
  structure of, 125*d*
Basic fuchsin, 2*s*
Basophilic substance, 18
BCOADH, branched chain 2-oxoacid dehydrogenase, 432
Beans, 419
Becker muscular dystrophy, 131
Beer-Lambert law:
  equation in studies of electron transport chain, 415*d*
  law, relationship, 69
  relationship, 53, 56
Bell-shaped curve, 262*d*
Benzene, structure, 153*s*
Bi Bi enzymatic reaction, 259*d*, 289
Bicarbonate:
  dehydration, 228
  in pyruvate carboxylase reaction, 353*d*
Biconcave disk, shape, 15, 16*d*
Bidirectional replication, 460
Bifunctional enzyme, 438, 440
Bilayer, lipid, 167*d*
Bile:
  acid definition, 168*s*
  cannaliculus, 17
  definition, 168*s*
  micelle, 363*d*
  salt, 387, 390, 395
  synthesis from cholesterol, 391*d*
Bilirubin, 453*s*
  diglucuronide, 453
  sulfate, 453
Binding energy, 234
Binding function, Adair, 269
Binding proteins, 267
Bioblast, 12
Biochemical standard state, 293
Bioenergetics, definition, 402
Biopterin, 455
Biosynthesis of purine nucleotides, 440
Biotin, 374*s*
Biotinylenzyme, 374*d*
Bird, 455
1,3-Bisphosphoglycerate, in substrate level phosphorylation, 402*s*
2,3-Bisphosphoglycerate, 344
  Rapoport-Luebering shunt and, 344
  synthesis, 344
Blood group substance, 16
Blood vessel, 17
Blue-green algae, 419
Boat, 36, 37*s*
Bohr, C., 267
Boltzmann, L., 238
Bond:
  angle, 234
  length, 234
  strain, 234
Bone, 11*t*, 17, 121*t*, 122*t*
  collagen, 148
Bongkrekic acid, inhibitor of ATP translocase, 415
Boric acid, pKa, 59
Bottlebrush structure, 127, 129
BPG:
  as an allosteric effector, 120
  binding to hemoglobin, 120, 151
  2,3-bisphosphoglycerate, 119*s*, 120, 121, 147
  competition with CO_2, 120
  low affinity for oxyhemoglobin, 120
  synthesis from glucose, 120
Brain, 11*t*, 21
  lipids, 362
Branched metabolic pathways, 265*d*
Branching enzyme, amylo-(1,4 → 1,6)-transglycosylase, 328*d*
Briggs, G. E., 256
Briggs and Haldane, 256
  analysis, 263*d*
5-Bromouracil, 199*s*, 200*s*
2,3-dideoxy-β-D-ribose, 220*s*
Brønsted-Lowry, definition of acid and base, 58, 233
Brown fat, uncoupled oxidative phosphorylation, 417
Brush border, 136
Buffer:
  amino acids as, 61–65
  pH, 59
Bulk properties of matter, 290
Butyrylcysteinyl-polypeptide, 376*d*

Ca^{2+}, Calcium ions:
  effect of ions on soaps, 183*d*
  in muscle, 17
Calorimeter, 291*d*
cAMP dependent protein kinase, 399*d*
Camphor, 185*s*
Cancer, 458
Cannaliculus, bile, 17
Cap-binding proteins, 505
Capping of RNA, 497
Carbamate, of hemoglobin, 121*s*
Carbamoyl phosphate, 435*s*
Carbamoyl phosphate synthetase:
  compartmentation of, 24
  location, 13*t*
Carbamoyl phosphate synthetase I, ammonia-dependent, 438
Carbamoyl phosphate synthetase II, 435
  glutamine-dependent, 438
Carbocation, 3, 247
Carbohydrate, 2, 21, 25
  classification, 25
  dietary, 374
Carbon dioxide, CO_2, 15
  binding to hemoglobin, 121
  hydration, 228
Carbon dioxide production:
  citric acid cycle, 249*d*, 348, 357
  pentose phosphate pathway, 340*s*
  pyruvate dehydrogenase, 352*d*
Carbon monoxide, inhibition of electron transport chain, 406*d*
Carbonic acid, 228
Carbonic anhydrase, 228
  Zn^{2+} in, 229
Carbonium ion, 247
Carboxypeptidase, 427, 428*d*
Carboxypeptidase A, 103*d*, 236, 247, 248*d*
  active site, 248*d*
Carboxypeptidase A (*Cont.*):
  catalytic mechanism, 248*d*
  EC number, 248
  exopeptidase, 248
  hydrolase, 248
  inhibitor, 236
  pseudo substrate, 236
  transition-state analogs, 236
  Zn^{2+} in, 248*d*
Cardiac muscle, 17
Cardiolipin, 159*t*
Carnitine, 369, 370*d*
Carnitine palmitoyltransferase, 392
Carnitine-fatty acid acyltransferase, location, 13*t*
Carnosine, 68*s*
Carotene, 163*s*, 165
β-Carotene, 163*s*, 165
Carotenoids, definition, 165
Cartilage, 121*t*, 122*t*, 127, 128*d*, 129
Cartilage binding, 131*t*
Cascade:
  of enzyme activity, 338
  in glycogenolysis, 338*d*
Casein, 74
Catalase, 12, 452
  molecular weight, 78, 82*t*
  sedimentation coefficient, 82*t*
Catalysts, 228
  acid, 233
  general acid, 247
  general acid-base, 232, 247
Catalytic mechanism, 245
Catalytic rate, 234
Cathepsin, 11*t*
CDP-ethanolamine, 381*d*
Cell, 1, 18
  adhesion to fibronectin, 130*d*
  cycle, 142
  cycle phases, G_1, S, G_2 and M, 461, 462*d*
  division, 458
  membrane, 1
  mucus-secreting, 17
  multinucleated, 17
  sex, 14
  shape, 12
  site of actin, 134
  structural hierarchy, 22
  types, 21
Cell-surface antigen, 16
Cellobiose, 45
Cellular nucleus, 15
Cellulose, 47
Central nervous system, lipids, 362
Centrifugal fractionation, 19
  of cells, 19
Centrifugation:
  differential, 4
  isopycnic, 5
Centrifuge, 5
Centriole, 6*d*, 12, 140, 143
  in cell, 142*d*
  in cilia, 142*d*
Centromere, 15, 15*d*, 142*d*, 143, 498
Centrosome, 140
Ceramidase, 381, 383*d*
Ceramide, 161*s*, 381, 382*d*, 383
Cereal protein, 419

Cerebronic acid, 161
Cerebroside, 381, 382
CETP, cholesteryl ester transfer protein, 365
Cetyl alcohol, structure, 153*s*
Chain termination triplets, 491
β-Chains, hemoglobin, 18
Chair, 36, 37*s*
Chair conformation, 247
Changeux, J., 270
Channel, 180*d*
Channeling of substrate(s), 442
Chaperonins, 87
Chemical equilibrium, 252
Chemical modification:
  carboxypeptidase A, 239
  of enzyme, 239
Chemical potential, 307
Chemiosmotic model, Mitchell's model of oxidative phosphorylation, 408
Chemiosmotic theory, 307
Chenodeoxycholic acid, 168, 391*d*
Chief cell, 226, 427*d*
Chiral center, 26, 29
Chitin, 47, 51
Chitosan, poly(α-δ-glucosamine), 51
Chloramphenicol, 507*s*, 516
Chlorophyll, 12
Cholate, 395
Cholecalciferol, 399
Cholecystokinin, 427
Cholesterol, 366, 368*s*, 391*d*, 392, 393*d*, 399
  acetyl-CoA and cholesterol biosynthesis, 387, 388*d*
  bile salts from, 387, 391
  metabolism, 387
  in mixed micelles, 167
  regulation of synthesis, 390
  structure, 164*s*
  synthesis from mevalonic acid, 182*s*
  vitamin D_3 from, 387
Cholesterol 7-α-hydroxylase, regulation of, 392
Cholesteryl ester, 365*d*
Cholesteryl ester hydrolase, 390
Cholesteryl ester transfer protein (CETP), 365
Cholic acid, 391*d*
  stereochemistry, 165*s*
  structure, 164*s*
Choline, 379*d*, 381, 397
  in phosphatidylcholine, 159*t*
Choline kinase, 380*d*
Choline phosphate, 379*d*, 381
Chondroitin sulfate, 47, 127, 128*d*, 131*d*, 149
Chondronectin, 131*t*
Chorion (egg) proteins, 463
Chorismate, 244*s*
Chorismate mutase–prephenate dehydrogenase, 244, 245
  adamantane in, 245
  inhibitor, 245
  transition state in, 245
  transition-state analog, 245
Chromatin, 14*d*, 216, 509
Chromatography:
  in amino acid analysis, 65, 66*d*
Chromatography (*Cont.*):
  ion exchange, 65, 66*d*
Chromomere, 15
Chromosome, 12, 14, 15*d*, 17, 24, 108, 142, 143*d*, 215, 458
  autosomes, 15
  homologous, 14
  in mitosis, 143*d*
  morphology, 15*d*
  number, 14
  segregation, 12
  sex, 15
  Y, 15
Chrondroitin sulfate, 11*t*
Chrysin, structure, 153*s*
Chylomicron, 168, 169*t*, 364*d*, 365, 379
  secretion of, 379*d*
Chymotrypsin, 80*d*, 239, 240, 248, 427
  α-, 428, 429*d*
  active site, 240
  δ-, 428, 429*d*
  enzyme-modification, 240
  phosphorylserine in, 239
  π-, 428, 429*d*
  serine in, 240
Chymotrypsinogen, 427, 428*d*, 429*d*
  molecular weight, 78, 83*t*
Cilia, 17, 108, 142*d*
  axoneme, 141, 142*d*
  centrioles, 142*d*
  in fallopian tube, 141
  inner sheath, 141, 142*d*
  nexin, 141, 142*d*
  plasma membrane, 141, 142*d*
  radial spoke, 141, 142*d*
Ciliated cell, 17
Circular DNA, 459
Circular duplex DNA, 224
Cistron, 494
Citrate, 374*ds*, 375*d*
  in the citric acid cycle, 346*s*, 350*s*
Citrate lyase, 374*d*
Citrate synthase:
  high substrate inhibition of, 350
  location, 13*t*
  NAD^+/NADH inhibition, 351
  reaction, 346*d*
  succinyl CoA inhibition, 351
Citric acid cycle, Krebs cycle, tricarboxylic acid cycle, 9, 311
  anaplerotic reactions of, 355*t*
  carbon dioxide production in, 357
  definition, 345, 346
  energetics of, 349
  location of, 9
  overall reaction, 357*d*
  reactions of, 346–349, 357
  regulation of, 350, 351
  studies of, using radioactive labels, 358
Citroyl-CoA, 350*s*
Citrulline, 433*d*
Classification of enzymes, 230*t*, 231*t*
Classification of hydrolases, 231*t*
Cleland convention, 289
Cleland, W. W., 259, 260
Closed system, definition, 290
Clostridium perfringens, 227
CoASH, coenzyme A, 370*d*
Coat protein of TMV, 109*d*
Cobalamin, vitamin B_{12}, 449
Cobalamin–N^5-methyl–THF: homocysteine methyltransferase, 449*d*
Cobalt, in vitamin B_{12}, 449
Cockroach(es), 420
Coding strand, sense strand, 489
Codon, 489, 490*t*
  anticodon pairing, 218
  glycine, 149
  initiation, 491
  nonsense, 491
  stop, 491, 504
  synonyms, 501
  termination, 491
Coenzyme A (CoA), 345*s*
  in pyruvate dehydrogenase, 116
Coenzyme Q, ubiquinone, 403, 404*s*
Coiled-coil, 151
  intermediate filament, 144*d*
  myosin, 137
  spectrin, 150
Colchicine, 150
Collagen, 121*t*, 123, 148, 149
  binding, 131*t*
  binding to fibronectin, 130*d*
  bone, 148
  cross-linking, 125*d*
  disulfide bonds in, 122
  exons in genes for, 149
  glycosylation of, 123
  hydrogen bonds in, 123
  hydrophobic, residues in, 148
  hydroxylysine in, 122
  hydroxyproline in, 122, 148
  introns in genes for, 149
  procollagen, 122, 123*d*
  prolylhydroxylase, 123
  properties, 122*t*
  protomeric, 149
  skin, 124
  tendon, 124
  triple helix, 93, 122, 123*d*, 125, 148, 149
  type IV, 125
  types, 122*t*
Collagenase, mammalian, 11*t*, 148
Compartmentation of metabolic pathways, 303
Competitive inhibition, 254, 255*t*, 262, 263*d*, 280*d*, 289
Complementary base pairs, 458
  DNA, 207*s*
Concentration gradient, 363
Concerted transition, 29, 30, 270
Condensation reaction, in sugars, 44
Configuration of atoms, 53*d*
  absolute, 26
  axial, 36*s*
  eclipsed, 36, 39
  energy difference, 36
  equatorial, 36*s*
  staggered, 36
Conformational change, 266
  actomyosin, 139*d*
  protein, 115
Conformational effect, 270
Conformer, 36
  boat, 37*s*
  chair, C1, 1C, 37*s*, 38*s*

Conformer (*Cont.*):
envelope, 39*s*
glucose, 52
twist, 39*s*
Conjugate base, 59
Connective tissue, 121, 129, 149
Consensus sequence:
DNA, 492
in mRNA, 497
Conservation of mass, law, 261
Contractile protein, 17
Contractile ring in cells, 150
Control of:
glycogen degradation, 329
glycogen synthesis, 329, 330*d*
glycolysis, 320, 334
pyruvate dehydrogenase, 352, 354*d*
Control enzyme, 265
Control points, of glycolysis, 317
Control signal, in metabolism, 299
Control strength of an enzyme, 310*d*
Conversion of, pyruvate to ethanol, 322*s*
Cooperative effect, hemoglobin, 147, 148
Cooperativity, hemoglobin, 147
Copper, electrochemical cell, 296
Copy number of genes, 460
CoQ-cytochrome *c* oxidoreductase, 405*t*
Core enzyme of DNA polymerase III, 466
Core protein:
in glycosaminoglycan, 127*d*
hyaluronan, 126
in proteoglycan, 128*d*
Corn, high lysine, 419
Cornea, 122*t*, 127
Cortisol, 368*s*
Cory, C. and Cory, G., 327
Cory cycle, 326, 327*d*, 342
Corynebacterium diphtheriae, 507
COT analysis of DNA, 214, 215*d*
Coulomb's law, 84
Covalent catalysis, 30, 231, 232*d*, 248, 249
Covalent enzyme catalysis, 241*d*
Creatine:
kinase, 422
synthesis, 450*d*
Creatinine, 456*s*
Crick, F. H. C., 29, 207, 458
Cristae, 12
Critical concentration:
of actin, 132–134*d*
gelsolin effect on actin's, 135
of tubulin in mitotic spindle, 150
of tubulin monomers, 140
Critical micellar concentration (CMC), 167
Crossover theorem, 300
in studies of oxidative phosphorylation, 418
Crowding, in solutions, 114
CsCl, in density gradient centrifugation, 7
CTP-choline: 1,2-diacylglycerol cholinephosphotransferase, 380*d*
CTP: cholinephosphate cytidylyltransferase, 380*d*
Cyanide, 229
antidote, nitrite, 418
Cyanide (*Cont.*):
inhibition of electron transport chain, 406*d*
poisoning, nitrite as antidote, 418
α-Cyano-3-hydroxycinnamate, inhibitor of pyruvate transport, 184
Cyanogen bromide, reaction with proteins, 80*s*
Cyclic AMP, cAMP:
in control of glycogenolysis and synthesis, 336*s*, 337*d*
structure, 336*s*
Cyclohexane:
boat, 36*s*
chair, 36*s*
Cycloheximide, 507*s*
Cyclooxygenase, 385*d*, 398*d*
Cyclopentane, 384, 386
Cyclopropane fatty acid, 182*s*
Cysteine, 202*s*, 233, 423*t*
disulfide bond, 68
modification by *N*-ethylmaleimide, 69
oxidation of, 68*s*
preventing oxidation of, 69
reduction of, 68*s*
S-carboxymethyl derivative of, 69
structure of, 55*t*
Cytidine deaminase, 243*sd*, 244
5,6-Dihydrouridine, 243*s*
transition-state analog, 244
Cytidine diphosphocholine (CDP-choline), 379*d*
Cytochemistry, 3
Cytochrome:
a, 405*t*, 406*d*
a_3, 405*t*, 406*d*
b, 405*t*, 406*d*
c, 404, 405*t*, 506*d*
c_1, 404, 405*t*, 506*d*
definition, 404
in electron transport chain, 405*t*, 406*d*
Cytochrome b_5, 397
Cytochrome *c*, molecular weight, 78, 81
Cytochrome oxidase, 405*t*
Cytokeratins, 145*t*
Cytokinesis, 142
contractile ring in, 144
plasma membrane in, 144
Cytoplasm, 7*d*, 10*d*, 19, 20*d*
staining, 19
Cytoplasmic membrane, 21
Cytoplasmic membrane of macrophages, 21, 22*d*
Cytosine, 12, 198, 199*s*, 458
Cytoskeleton, 6*d*, 12, 130, 131
actin in, 137*d*
anion channel attachment to, 136, 137*d*
ankyrin, 136, 137*d*
protein 4.1, 136, 137*d*
proteins in, 136
spectrin in ,136, 137*d*, 149, 150
submembrane, 15

D loop in DNA, 461
dAMP, deoxyadenylic acid, 203
Dansyl chloride, 79*s*
Daughter DNA, 459
DCCD, dicyclohexyl carbodiimide, 412
DCCD, dicyclohexyl carbodiimide (*Cont.*):
inhibitor of oxidative phosphorylation, 412
ddATP, 474
2′, 3′-ddCTP, 486
ddNTP, 474
DEAE-cellulose, 78*s*, 102
Debye-Hückel theory, 305
Deformylase, 505
Degenerate code, 489
Delft, in Holland, 1
Denaturation of DNA, 479, 480*d*
Denaturation temperature, DNA, 12
Density-gradient centrifugation, 5
2-Deoxy-β-D-ribose 39, 40*s*, 201*s*
Deoxy sugars, 39
Deoxyadenosine, 221*s*
Deoxyadenylic acid, dAMP, 203
Deoxycholic acid, 168
Deoxycytidine, 202
Deoxyguanosine, 202*s*
Deoxyhemoglobin, 119, 121, 148
cooperative effects, 119
oxygen affinity, 119
sigmoidal oxygen dissociation curve, 119
tense structure, 119
Deoxynucleoside triphosphate (dNTP), 465
Deoxyribonuclease (DNase), 219, 446, 447*d*
Deoxyribonucleoside, 201
Deoxyribonucleotide, 202
Deoxyribose, 198
Deoxyribose 5′-phosphate, 488*s*
Deoxyribose nucleic acid, DNA, 198
Deoxythymidine, 202
Deoxythymidine 5′-phosphate, 202*s*
Depot fat, triacylglyerol in, 158
Derivation of rate equation:
algorithmic method, 257
King and Altman, 257
reaction patterns, 257, 258*d*
Desaturation, of fatty acid, 376
Desmin, 145*t*
Desmosine, 126
Desmosome, 145
Desolvation reaction, 234
Detergent, 363*s*
properties of bile salts, 168
Dextran, 78
Dextrin, 228
DFP (diisopropyl fluorophosphate), 239*s*
Diabetes mellitus, 344
Diacylglycerol, 158
Diad symmetry, 111
Dialkyl phosphoamidite (DPA), 211
Dialysis, in vesicle preparation, 170
Diastase, 228
Dicarboxylate carrier, in mitochondria, 324
Dicarboxylate system, of amino acid transport, 431
1,2-Diacylglycerol, 378*d*, 380*d*, 382*d*
Dicyclohexyl carbodiimide (DCCD), inhibitor of oxidative phosphorylation, 16*d*, 412
Dideoxy method of Sanger, 473

2′,3′-Dideoxynucleosides, 473
Dielectric constant, 85
Dietary carbohydrate, 374
Dietary lipid, 391
Differential centrifugation, 4, 5*d*, 24, 303, 308
Differentiation, 1
Diffusion:
  coefficient, 81
  lateral, in membranes, 173
  rotational, in membranes, 173
Digestion, 228
  of dietary lipid, 391
  of lipids, 362, 363
  role of bile salts in, 391
  of triacyl glycerol, 363*sd*
Digestive enzyme, carboxypetidase A, 247
Dihedral angle, 111
Dihydrobiopterin, 425*d*, 426*s*
Dihydrobiopterin reductase, 426*d*, 455
Dihydrofolate, dihydrofolic acid, 448*s*
Dihydrofolate reductase, 443, 444*d*
Dihydrolipoyl dehydrogenase (E_3), in pyruvate dehydrogenase, 116, 352, 353*d*
Dihydrolipoyl transacetylase, 352
Dihydrolipoyl transacylase (E_2), reaction, 352, 353*d*
Dihydroorotate, 437
Dihydroorotate dehydrogenase, 438
Dihydroorotate synthetase (or CAD), 438
Dihydroxyacetone, 30, 31
Dihydroxyacetone phosphate, in glycolysis, 314*s*
20, 22-Dihydroxycholesterol, 393*d*
3,4-Dihydroxyphenylalanine (dopa), 432*s*
Diisopropyl fluorophosphate (DFP), 239*s*
Dilution principle, Ostwald's, 147
Dimensional analysis, 252
Dimerization:
  association constant, 114
  of a protein, 146
Dimethoxytrityl group (DMT), 211
Dimethylallyl pyrophosphate, 389, 390*d*
Dimethylallyl transferase, 390*d*
Dinitrophenol, effect on oxidative phosphorylation, 412, 416
Dipeptidase, 430
Diphosphatidylglycerol, 159*t*
Diphtheria toxin, 507
Diploid, 14
  cell, 14
Disaccharide, 25, 44, 126, 127*d*, 247,
Discontinuous DNA replication, 467
Disease, inherited, 21
Dissociation equilibrium constant, 255
Disulfide bond, bridge, 80, 240*s*
Disulfide bond (S—S) rearrangease, 241
Dithiothreitol, reduction of disulfide bonds by, 68*s*
DNA, 7, 11*t*, 12, 14, 17, 19
  amplification, 87, 88*d*, 464*d*
  annealing, 87, 99, 212, 479
  B form, 209
  base composition in various species, 206*t*
  base pairing, 207
DNA (*Cont.*):
  base-pairing geometry, 207*s*, 209
  buckle, 209
  chain polarity, 205
  circular, 459
  cloning, 238
  complementary, 207
  COT analysis, 214
  denaturation, 212*d*
  denaturation temperature, 12
  deoxyribose in, 198
  deoxyribose nucleic acid, 7
  double helix, 468
  duplex, 459
  excision repair of, 475
  eye form, 460
  formamide effect on, 212
  helix, 458
  highly repetitive, 498
  hyperchromic effect, 212
  major groove, 208, 209*d*
  melting temperature, 214*d*
  minor groove, 208, 209*d*
  mitochondrial, 12, 461
  moderately repetitive, 498
  mutation, 458
  negative supercoiling, 468
  nick in, 464
  nonhomologous recombination, 463
  nonrepetitive, 498
  in nucleus, 14
  Okasaki fragments, 222
  opening, 209
  photoreactivation of, 476
  pitch, 209*t*
  positive supercoiling of, 468
  primers, 479
  propeller, 209
  repair, 458
  repair synthesis, 458
  replication, 458
  replication fork, 459
  rise, 208, 209*t*
  rolling-circle mechanism of replication, 464*d*
  satellites, 15
  semiconservative replication, 459
  separation by centrifugation, 7
  sequence heterogeneity, 214
  size, 17
  staining, 2, 19
  structure, 206*d*
  as a substrate, 11*t*
  template, 481
  termination of replication, 460
  thymine dimer in, 475*s*
  topology, 215, 459*d*
  unstacking, 212
  uv damage, 458
  wild type, 238
  X-ray diffraction, 208
  Z-form, 208, 210*d*
DNA gyrase, 468, 469*d*, 470
DNA helicase, 468, 469*d*
DNA ligase, 467, 468*d*, 469*d*, 476*d*, 477, 478*d*
  NAD^+, reactant with, 468
DNA mutation, 470
DNA polymerase, 465*d*
DNA polymerase (*Cont.*):
  I, 465, 469*d*, 474*d*, 476*d*, 483
  II, 465
  III, 465, 483
DNA polymerase α, 472
DNA polymerase δ, 472
DNA polymerase III, holoenzyme, 466, 467, 469*d*, 470
DNA primase, 485
DNA synthesis, 458
DnaA, 470
  boxes, 470
DnaB, 470
  helicase, 468, 469*d*
DnaC, 470
DnaG, 470
Dnase, deoxyribonuclease, 219
Dolichol, 163*s*
Domain of protein, 106
Dopa quinone, 432*s*
Double-helical DNA, 458
Double helix, actin, 113, 132
Double-reciprocal plot, 253*d*, 264*d*, 266*d*
Down syndrome, 15, 24
  chromosomal translocations in, 24
  facial features, 15
Drosophila melanogaster, 463
Drug resistance in cancer, 445
dTTP, 465
Duchenne muscular dystrophy, 131
Duclaux, 228
Dye, 2
  supravital, 23
Dynein, 141*d*, 144
  arms, 142*d*
  in family of motors, 143

Eadie and Hofstee, plot, 288
EC group, 240, 245
EC integer, number, 229, 230*t*
  arginase, 230
EcoRI, 226, 477*d*, 478*d*
Edman degradation, 79*d*
EF-G, elongation factor G, 504
EF2, elongation factor, 508
Effector, 266
Egg albumin, 77
Ehlers-Danlos syndrome VII, 149
Elaidic acid, 156
Elastase, 427, 428*d*
Elastin, 121, 125, 231
  extracellular matrix, 125
  fibers, 126
  hydrophobic interactions in, 126
  lubrication of, 126
  oiled coil, 126
  reverse turns, 125
Electrochemical cell, 296, 297*d*
Electrochemical potential, definition, 307*d*
Electrochromism, 407
Electromagnetic radiation, 3
Electromotive force (E), 296*d*
Electron, 3
  depletion, 232
  withdrawal, 232
Electron microscopist, 21
Electron microscopy 3, 4, 23
  autoradiography, 23

Electron microscopy (*Cont.*):
of macrophage, 21, 22*d*
Electron spin resonance, 4-dinitrophenol and, 418
Electron transport chain:
components, 402
CoQ-cytochrome *c* oxidoreductase, complex III, 405*t*
cytochrome oxidase, complex IV, 405*t*
definition, 402
inhibitors, 406*d*
NADH/CoQ oxidoreductase, complex I, 405*t*
organization of, 405*t*
respiratory chain, 402
succinate/CoQ oxidoreductase, complex II, 405*t*
Electronegativities of atoms in proteins, 85
Electrophilic attack, 231, 232
Electrophoresis:
anode, 77
cathode, 77
Electrophoretic gel, 2
Electrostatic interactions, 123
in proteins, 84
Elementary chemical process, definition, 251, 252
eIF-2, 452
Elongation of DNA, 460
Elongation factor:
G (EF-G), 504
Tu, 503
Enantiomer, 26
Endocytosis, 9, 10*d*
Endonuclease, 219
Endopeptidase, 248
Endoperoxide, 386
Endoplasmic reticulum (ER), 6*d*, 9*d*, 16, 19, 20*d*, 138
in adipocyte, 16*d*
rough (RER) 4, 7*d*
smooth (SER), 4
Energetics of glycolysis, 316
Energy barrier, 237
Energy change, 293
Enhancement rate, of enzymes, 230
Enhancer, 494
Enolase, 316*d*
Enoyl reductase, 377*t*
Enoyl-CoA hydratase, 372*t*
Enteropeptidase, 427, 428*d*
Enthalpy (*H*), 114, 238, 290, 292
Enthalpy change, ΔH, 292
Entropy (*S*), 114, 234, 238, 290
Entropy change, ΔS, 292
Enzyme(s), 7, 11*t*, 18, 251
activator, 228, 254
active site, 230
acylation by, 231
catalysis, 228
chemical modification, 238
classification, 230*t*, 231*t*, 240
cofactor, 229
cofactors, 251
crystallization, 228
digestive, 247
effective concentration of substrates, 230
Enzyme(s) (*Cont.*):
effectors, 228
enhancement rates, 230
glycosylation by, 231
group-transfer by, 231
inhibition, 254
inhibitors, 228, 243, 244
mechanisms, 254, 259, 260
metal ion in, 229
metal ions, 251
models, 254
multimeric form, 229
multiple-subunit complex(es), 229
optimum pH, 261
pH effect on, 281*t*, 282*d*
pH, effects on, 251
pH effects on rate, 261*d*, 262*d*
phosphorylation by, 231
propinquity, effects in, 230
prosthetic group, 229
proximity effects in, 230
rate enhancement, 228, 233, 237, 238, 241, 249
reactants, 228
reaction intermediate, 231
regulation, 238
specificity, 228, 238
stability, 238
structural variants, 228
substrates, 228
temperature, effects on, 251
tetrahedral intermediate, 232*d*, 238, 244
transition state, 231–235, 237, 243, 244
Enzyme Classes, 230*t*
Enzyme Commission (EC), 229, 241
Enzyme inhibition, 254, 255*t*, 263
anticompetitive, 254
competitive, 254
mixed, 254
noncompetitive, 254
Enzyme kinetics, 251
Enzyme mechanism, 256
Enzyme unit, of activity, 257
Enzyme-inhibitor complex, 263*d*
Enzyme-substrate complex, 234, 235, 255
Enzymology of DNA replication, 483
Eosin, 2*s*, 21
staining of mast cell, 21
Epidermolysis bullosa, 24
Epimer, 38, 49, 241
Epinephrine,
in glycogenolysis, 329, 432, 433*s*
Epithelia, epithelium, 17
Equatorial configuration, 36*s*
Equilibrium:
chemical, 228
constant, 114, 237, 252
definition, 255
relationship to ΔG, 293, 294
Erythritol, 52
D-Erythro-pentulose, 30
Erythrocyte, 15, 22
cytoskeleton, 136, 137*d*
red blood cell, 15, 16*d*, 119
D-Erythrose, 27, 28, 30
4-phosphate, 341*s*
L-, 28
Escherichia coli (*E. coli*):
chromosome size, 224
genome size, 215, 216*t*
membrane composition, 172
Eskimos, Inuit, 419
Essential amino acid, 419
Essential fatty acid, 376
Ester bond, 231*t*
Ester of fatty acid, 11*t*
Estrogen, 393*d*
Ethanol, 397
Ethanolamine:
in phosphatidylethanolamine, 159*t*
in phospholipid synthesis, 381*d*
Ether bond, 231*t*
Ethidium, 473
Ethylene glycol, 157*s*
1,2-Propanediol, 157*s*
N-Ethylmaleimide, 69*s*
effect on oxidative phosphorylation, 417
modification of cysteine by, 69*s*
Euchromatin, 509
Eukaryote, genome size, 216*t*
Europeans, 419
Evolutionary tree, of proteins, 98*t*
Excision repair, 476*d*
of DNA, 475, 486
Exergonic reaction, 237, 419
Exocrine (secretory) gland, 8
Exon, 497
Exonuclease, 219
5′ → 3′Exonuclease, 465, 466
activity of DNA polymerase I, 466
Exopeptidase, 236
Extension process, in PCR, 480*d*
Extracellular, 21
composition, 121*t*
matrix, 108, 129, 130
proteoglycans in, 126
space, 10*d*, 20*d*
Extrachromosomal DNA, 463
Extrinsic binding constant, definition, 269
Extrinsic parameters, 271
Eye form of DNA, 460
Eyring, H., 237, 238

F, Faraday constant, 296*d*, 407*d*
F_0, of ATP synthase, 412
F_1, of ATP synthase, 412
F-actin, 134*d*
Fabry disease, 383*d*
Facial features, in Down syndrome, 15
Facilitated diffusion, across membranes, 177, 178*d*
FAD, flavin adenine dinucleotide, 403
in pyruvate dehydrogenase, 116
structure, 312
$FADH_2$, in fatty acid oxidation, 370
Fallopian tube, cilia, 141
Familial hypercholesterolemia, 367, 368
Faraday constant (*F*), 296*d*, 407*d*
Farber disease, 383*d*
Farnesol, 162*s*, 185*s*
Farnesyl:
pyrophosphate, 390*d*
transferase, 390*d*
Fast axonal transport, of vesicles, 140, 141*d*

Fat tissue, 16
Fat vacuole, in adipocyte, 16*d*
Fatty acid, 116, 117, 352*s*, 363*s*, 364, 378, 381, 392
  acylate, 155
  branched, 135, 157*s*
  cis and *trans* double bonds, 135, 156
  cyclopropane containing, 182*s*
  ester, 155
  melting point, 156
  number notation, 156*s*
  polyunsaturated, 155*s*
  saturated, 155*s*
  structure, 154*s*
  synthesis of, 374
  transport into mitochondria, 369
  unsaturated, 155*s*
Fatty acid activation, 369
Fatty acid–CoA ligase, location, 13*t*
Fatty acid synthase, enzyme complex, 116, 376, 396
Fatty liver, 397
FDNB, fluoro-2,4-dinitrobenzene, 102
Feathers, 106
Feedback control, of enzyme, 283*d*
Feedback inhibition, of enzyme, 265*d*
Fermentation, 228
Ferritin:
  molecular weight, 83*t*
  self association, 111
Ferrochelatase, 451
Fetal hemoglobin, 147
Fetus, 15
Fibrin, binding to fibronectin, 130*d*
Fibroblast, 121, 130
  binding, 131*t*
  desmosomes in, 145
  and thymosin, 134
Fibronectin, 121, 131*t*
  as cold-insoluble globulin, 130
  proteolytic cleavage of, 130
  structure, 130*d*
Field, J. W., 1
Filmer, 273
Finger nail, 145
First committed step, 308
First messenger:
  epinephrine, 337
  glucagon, 337
First-order reaction, 252, 256*t*, 274
Fischer projection formula, 26
Fischer projection of triacylglyerol, 157*s*
Fischer's lock and key concept, 273
Fish, 455
Fixation:
  glutaraldehyde, 23
  in histology, 1, 23
Fixed nitrogen, 419
Flagella, flagellum, 108, 141, 142*d*
  sperm, 141
Flavin adenine dinucleotide (FAD), 403
Flavin mononucleotide (FMN), 403
Flip-flop movement, phospholipid movement across membranes, 173
Fluid mosaic model of membrane, Singer and Nicolson, 172, 173*d*
Fluoro-2,4-dinitrobenzene (FDNB), 79*s*, 102
Fluoroacetate, '1080', 359, 360*s*
Fluorodeoxyuridylate, 473
5-Fluorouracil, 219, 220*s*, 444*s*, 445*d*
Flux, definition, 252, 301
Flux control coefficient, 310*d*
FMN, flavin mononucleotide, 403
Folate transporter, 444
Folic acid, 424
  folate, derivatives, 447, 448*s*
Folylpolyglutamyl synthetase, definition, 444
Foreign body, in macrophage, 22*d*
Formate, 448
Formation of cholesterol, 391*d*
N^{10}-Formyl tetrahydrofolate (10-CHO-THF) synthetase, 442
Formylation of methionyl-tRNA, 503, 504*d*
Fractional saturation, 271, 286
Free energy, G, 237
  of binding, 234
  Gibbs, 290, 293
  of reaction, 235*d*
Free-radical, of ubiquinone, 404
Freeze-clamping, 299
Frequency of initiation, in DNA replication, 463
Frog, 455
Fructokinase, 330, 331*d*
Fructose, 51
  in gluconeogenesis, 330, 331*s*
Fructose-1,6-bisphosphate, 242, 314*s*
α-D-Fructose 1,6-bisphosphate, 313*s*
  in gluconeogenesis, 325*s*
  in glycolysis, 313*s*
Fructose 2,6-bisphosphate, control of in glycolysis, 301*d*
Fructose-1,6-bisphosphate aldolase, 242
  in glycolysis, 314*d*
Fructose-1,6-bisphosphatase, 301*d*, 314*d*
  in gluconeogenesis, 325*d*
Fructose-1-phosphate aldolase, 331*d*
α-D-Fructose 6-phosphate:
  in gluconeogenesis, 325*s*
  in glycolysis, 313*s*
  in the pentose phosphate pathway, 341*s*
Fructoside, 43
L-Fucose, 40*s*
Fumarase:
  fumarate hydratase, 13*t*, 249, 348*d*
  location, 13*t*, 436*s*
Fumarate, in the citric acid cycle, 348*s*
Fumaric acid, oxidation energetics, 304
Furan, 35*s*
Furanose, 49
Fusidic acid, 507

ΔG, change in Gibbs free energy, 293
G-actin, 134*d*
G_1 phase of cell cycle, 461, 462*d*, 463
G_2 phase of cell cycle, 461, 462*d*, 463
GABA, γ-amino butyrate, 455
Galactitol, 52
Galactokinase, 330, 331*d*
α-D-Galactopyranose, 38*s*
Galactosamine D-(2-amino-2-deoxy-D-galactose), 42
D-Galactose, 28*s*, 47, 49, 51
  optical activity, 52
  optical rotation, 52
Galactose hydroxylysine, 123
Galactose-1-phosphate uridylyltransferase, 330
α-Galactosidase, 383*d*
β-Galactosidase, 11*t*, 383*d*
  in *E. coli*, 508
Galactoside, 11*t*
Galactoside permease, in *E. coli*, 508
Galactosylceramide, 161*s*
Galactosylceramide-β-galactosylhydrolase, 383*d*
Galactosyltransferase, 149
Galacturonic acid, 47
Gamete, 14
Ganglioside, 381, 382, 383
  definition, 161
  GM_1, 383*d*
  GM_2, 383*d*
  GM_3, 383*d*
Gastric gland, 427*d*
Gastrin, 426
GATA-1, in erythroid cells, 496
Gaucher disease, 21, 382*t*, 383*d*
Gel electrophoresis, for determining protein molecular weight, 83
Gel filtration, 102
  molecular weight estimation, 82*t*, 83*dt*
  of proteins, 78, 82, 83*d*, 83*t*
Gel-to-sol transition, actin, 135
Gelatin, 70, 121
Gelsolin, 135
Gene, 458
Gene amplification, 444, 463
Gene cloning, 477
Gene cluster, 499*d*
Gene or DNA amplification, 458
General acid catalysis, 247
General acid-base catalysis, 247
Generation-time of bacteria, 462
Genes, isolation of, 476
Genetic code, 489, 490*t*, 509
Genome, 14, 458
Geraniol, 162*s*
Geranyl pyrophosphate, 390*d*
Geranyl transferase, 390*d*
Gibbs free energy, 238, 290, 293
  change, ΔG, 293
Gibbs, W., 238
Glial filaments, 145*t*
Globin:
  gene cluster, 499*d*
  synthesis, 509, 516
Globin fold:
  in hemoglobin, 117, 147
β-Globin gene, 496, 497*d*
Globoside, 162*s*
Glucagon, in glycogenolysis, 329
Glucan, 45
Glucitol, 51
  D-, 40*s*, 41*s*
Glucocorticoid, 393*d*
Glucofuranose, α-, β-, 35
Glucogenic amino acids, 432
Glucokinase:
  in liver, 317
  K_m for glucose, 317, 318*d*
Glucomannan, 45
Gluconeogenesis:
  definition, 323

Gluconeogenesis (*Cont.*):
scheme, 324*d*
D-Gluconic acid, 41*s*, 47, 51, 52
δ-Gluconolactone, 41*s*
α-D-Glucopyranose, 35, 38*s*
D-Glucosamine, 2-amino-2-deoxy-D-glucose, 42, 51
D-Glucose, 3, 19, 28*s*, 29*s*, 32*s*, 33, 47, 123, 396
anomeric carbon, 33
anomers α-, β-, 32, 34*s*, 35*s*, 48
conformation 35, 48
Fischer projection, 32*s*, 34
in glycolysis, 311
Haworth projection formulas, 32, 33, 37*t*, 48*s*
hexokinase and, 237*s*
ionophore effects on transport, 186
NMR of, 48
open or straight chain structure, 32
optical rotation, 48
product of phosphatase reaction, 19*d*
reducing carbon, 33
transport across membranes, 175, 183, 184
transport modes, 179*t*
α-D-Glucose:
in gluconeogensis, 326*s*
in hexokinase, 313*s*
Glucose 1-phosphate, in glycogen synthesis, 327*s*
Glucose 6-phosphate, 19, 249*t*, 317*s*
in glycogen synthesis, 327*s*
hexokinase and, 237*s*
in the pentose phosphate pathway, 339*s*
Glucose-6-phosphate isomerase, second step in glycolysis, 313*d*
α-D-Glucose 6-phosphate
in gluconeogensis, 326*s*
in hexokinase, 313*s*
Glucose-6-phosphatase, 19, 19*t*, 249*t*
in gluconeogenesis, 326*d*
staining, 19*d*
α-Glucosidase, 11*t*
β-Glucosidase, 11*t*, 21, 383*d*
Glucoside, 43
N-Glucoside, 43*s*, 201*s*
Glucuronate transferase, 149
Glucuronic acid, 126
D-Glucuronic acid, 52
β-Glucuronidase, 11*t*
Glutamate, 420*s*, 421*s*, 422*s*, 423*t*
as pH buffer, 72
structure of, 55*t*
L-Glutamate, 240*s*
Glutamate dehydrogenase, 420, 423*d*, 434
reaction, 453*d*
Glutamate synthetase, 420*d*
Glutamic acid, 233, 247
isoelectric point, 64*t*
pH titration of, 63*d*
Glutamine, 420*s*, 421*s*, 423*t*
structure of, 54*t*
Glutamine synthetase, 420*d*
thermodynamics, 299*d*
γ-L-Glutamyl-L-cysteine, 240*d*
Glutaraldehyde fixation, 23
Glutathione, 67*s*, 240*d*, 241
Glutathione synthetase, 241
Glycan, 45
Glyceraldehyde, 30
D-, 26, 27
L-, 26–28
Glyceraldehyde kinase, 331
Glyceraldehyde 3-phosphate, 240, 250
in glycolysis, 314*s*
in the pentose phosphate pathway, 341*s*
in substrate level phosphorylation, 402*s*
Glyceraldehyde-3-phosphate dehydrogenase, 241, 315*d*
active site, 250
acyl compound, 250
and 1, 3-bisphosphoglycerate in glycolysis, 315*s*
glyceraldehyde 3-phosphate in, 250
Glycerol, 363
in diphosphatidylglycerol, 159*t*
in gluconeogenesis, 331, 332*s*
in glycerolipids, 157*s*
in phosphatidylglycerol, 159*t*
Glycerol kinase, 331, 332*d*
Glycerol 3-phosphate, 331
Glycerol-3-phosphate dehydrogenase:
cytoplasmic, 333, 334*d*
in glycerol 3-phosphate shuttle, 333, 334*d*
mitochondrial, 333, 334*d*
Glycerol 3-phosphate shuttle, 333, 334*d*
Glycerolipids, 154, 157
Glycerose, 29
Glycine, 67*s*, 238, 240*d*, 395, 423*t*, 424*s*
in bile salts, 168*s*
codon, 149
pH titration of, 61*d*
structure of, 54*t*
synthesis, 424*d*
Glycocholate, 168*s*, 395
Glycogen, 11*t*, 17, 23, 46
granules, 327
site of synthesis, 17
structure, 328*s*, 329*d*
synthesis of, 327
Glycogen phosphorylase, 329*d*
Glycogen synthase, 328*d*
Glycogenolysis, control of, 329, 330*d*
Glycoglycerolipid, structure, 160*s*
Glycolipid, 381
Glycolysis:
control points of, 317
definition, 311
effects of hormones on, 336
energetics of, 316, 317*d*
in erythrocytes, 15
Glycolytic enzyme, 242
Glycolytic flux, 335
Glycolytic pathway, 15
Glycoprotein, 1, 2
Glycosaminoglycan, 47, 122, 126, 127*d*, 129*d*, 130
binding, 131*t*
chondroitin sulfate, 127*d*
Glycoside, 43
O-Glycoside, 43*s*, 44
Glycosidic bond, 25, 42, 45, 50, 245
β-Glycosidic linkage, in nucleic acid, 201*s*
Glycosphingolipid, 11*t*, 21, 161, 382*d*
Glycosyl bond, 231*t*
Glycosylation, 43, 231
of collagen, 123
Glycosyltransferase, 126, 149
Glycylalanine, 67*s*, 89*s*
Glycylglycylglycine, 75
Glyoxylate, in the glyoxylase cycle, 356*s*
Glyoxylate cycle:
overall reaction, 357*d*
reactants, 355, 356*d*
Glyoxysome, 356
Golgi apparatus, 4, 8*d*, 9, 9*d*, 17, 20*d*, 22*d*, 155*s*, 156*s*, 506
body, 6*d*
body size, 17
in macrophage, 22*d*
Gomori, reaction, 3, 4
Gout, 456
Gramicidin A, 181
Granules, Palade's, 13
Granulocytes, 130
Graphical evaluations, in enzyme kinetics, 276
Graphical procedures, kinetic analysis, 254
Group translocation, in transport, 179*d*
Growth factor, 129
GTP, 140, 504*d*
caps on tubulin, 140
GTP/GDP, in the citric acid cycle, 347*d*
Guanidine hydrochloride, 87*s*
Guanidinoacetate, synthesis, 450*d*
Guanine, 200*s*, 458
deaminase, 447*d*
Guanosine, 12, 202
Gulose, 28*s*
Gut lumen, 17

H, enthalpy, 290, 292
Δ*H*, enthalpy change, 292
H^+/phosphate symport, in mitochondria, 409
H^+/ATP ratio, in oxidative phosphorylation, 408*d*
H_2, in nitrogenase reaction, 420*d*
Hair, 106, 145
Haldane, J. B. S., 256
Half-cell:
electrode potential, 296, 297*d*
reaction, 295*d*
Halide bond, 231*t*
Halobacterium halobium, membrane composition, 172
Hanes and Woolf, plot, 288, 289
Haploid cell, 14
Haworth projection formula:
D-glucose, 32, 33, 37*t*, 48*s*
3,6-anhydro-D-glucose, 48, 49*s*
HDL, high-density lipoprotein, 168, 169*t*, 390
Heart, 17
Heat capacity, 292
Heat change, Δ*q*, 292
α-Helical, intermediate filament, 144*d*
α-Helical tail, myosin, 137
Helicase, 476*d*
DnaB, 469*d*
α-Helices, in hemoglobin, 117

α-Helix:
helix breaking amino acid residues, 91*t*
protein, 112
spectrin, 90, 91*d*, 150, 151
Helix:
A-form, 209*t*
B-form, 209*t*
buckle, 210*s*
DNA, 207, 458
double, 207
major groove, 209*d*
minor groove, 209*d*
opening, 210*s*
pitch, 107
propeller, 210*s*
right-handed, 207
roll, 209, 210*s*
tilt, 209, 210*s*
triple-stranded, collagen, 121
twist, 209, 210*s*
weight tested, 207
Z-form, 209*t*
Helix-destabilizing protein, DNA, 468
Hemagglutinin, trimer, 112, 113*d*
Heme, 147
in cytochromes, 404
group, 118*d*
synthases, 451
Hemiacetal, 32*s*, 33*s*, 43*s*
Hemicellulose, 47*d*
Hemin, 452, 509
Hemoglobin, 15, 17, 18, 22, 23, 147, 148, 151, 344
A. V. Hill and, 267
Adair binding model, 269
Adair equation and, 289
allosteric interactions in, 118, 119*d*
binding to, 267
2,3-bisphosphoglycerate binding, 267
Bohr effect, 120
BPG binding, 147, 151
β chain, 18
comparison with myoglobin, 151
cooperative effect, 147, 148
cyclic symmetry, 118*d*
degradation, 452
deoxyhemoglobin, 148
dihedral symmetry, 118*d*
electrostatic interactions, 118
fetal, 145
fractional saturation, 267
globin fold, 147
heterotropic effector, 267
Hill coefficient, 148, 151
Hill equation and, 289
hydrogen bonds, 118, 148
hydrophobic interactions, 118
molecular weight, 82*t*
MWC binding model and, 271
oxy-, 118
oxygen binding, 119*d*, 147, 148, 267
packing contacts, 117*d*, 118
proximal histidine, 118, 147
quaternary structure, 97, 118, 147
relaxed structure, 148
R_S value, 289
salt links, 119
sedimentation coefficient, 82*t*
sigmoidal binding curve, 267
Hemoglobin (*Cont.*):
. sliding contacts, 117*d*, 118, 148
tense structure, 121, 148
x-ray analysis, 97
Hemoglobin tetramer, 18
Hemolymph, insect, 52
Henderson-Hasselbalch equation, 60, 72, 73
derivation, 306*d*
Heparan sulfate binding, 131*t*
Heparin binding to fibronectin, 130*d*
Hepatocyte, 17
liver cell, 12, 17
Heptad repeat, 150
spectrin, 151
Hess's law, of constant heat summation, 304
Heterochromatin, 509
Heterocyclic compound, 2
Heterogeneous nuclear RNA (hnRNA), 497
Heterolytic bond cleavage, 245, 247
Heterophagy, 9, 10*d*, 19
Heterotropic effect, 267, 271, 273*d*
Heterotropic effector, definition, 266
Hexadecanoate, 369*s*
Hexadecanoyl-CoA, 369*s*
Hexagonal close-packing, of spheres, 113
Hexokinase, 232*d*, 236, 237, 276, 344
ATP as substrate, 236
binding energy of, 237
first step in glycolysis, 313*d*
glucose 6-phosphate product, 237*s*
glucose as substrate, 236*s*
induced fit in, 237
K_m for glucose, 317, 318*d*
nucleophilic attack in, 237
product inhibition of, 317
substrates, 237
Hexosaminidase A and B, 383*d*, 397
Hexose, 25
Hexose monophosphate, 339
Hierarchy of structure in cells, 17
High-density lipoprotein (HDL), definition, 168, 169*t*, 364, 390
High-lysine corn, 419
Hill, A. V., 151, 267, 268
Hill coefficient, 268, 287
in hemoglobin, 148, 151
Hill equation, 267, 286*d*, 289
Hill plot, 268
Histamine, 21, 22*sd*, 240*s*, 455
in mast cells, 21
Histidine, 233, 423*t*
as pH buffer, 72
structure of, 56*t*
L-Histidine, 240*s*
Histidine decarboxylase, 241
Histidyl, 231
staining, 2
Histochemistry, 3, 23
Histology, 2, 18
Histone, 14, 216, 473
genes for, 498
H1, 217*d*
H2A, 217*d*
H2B, 217*d*
H3, 217*d*
H4, 217*d*
HMG-CoA lyase, 373*t*
HMG-CoA reductase, 388*d*, 394, 399
HMG-CoA synthesis, kinase, 399*d*
hnRNA, heterogeneous nuclear RNA, 497
Hogness box, TATA box, 494, 496
Holland, 1
Homocysteine, 449*s*
transmethylase, 449
Homogenate, 4
Homogenization, 5
Homogenizer, 19
Homo-γ-linolenic acid, 398
Homologous chromosomes, 14
Homolytic bond cleavage, 245
Homoserine dehydrogenase, 283
Homotropic effect, 266, 267, 271, 273*d*, 270
Hormone-sensitive lipase, 368, 392, 394
Horse-liver alcohol dehydrogenase, 289
Human, cytochrome *c*, 98*td*
Hyaloplasm, 6*d*
Hyaluronan, 127*d*, 128*d*, 129
core protein, 126
Hyaluronic acid, 11*t*, 47
Hyaluronidase, 11*t*, 129
Hydride ion, 403
Hydrogen, in nitrogenase reaction, 420*d*
Hydrogen bond, 86*t*, 247
with amino acids, 75
between bile salt molecules, 168
in cellulose, 47
in collagen, 123
in hemoglobin, 148
Hydrogen electrode, 297*d*
Hydrogen peroxide, formation, 12
Hydrolase, 229, 230*t*
classification, 231*t*
Hydrolyic cleavage, 230*t*
Hydrolysis, 231*d*
Hydrolysis reaction, 233, 234
Hydrolytic enzyme, 16, 21
in erythrocytes, 16
Hydronium ion, 57*d*
5-Hydroperoxyeicosatetraenoic acid, 387*d*
Hydrophobic effect, with lipids, 166
Hydrophobic interaction, 86*t*, 123, 234
in elastin, 126
in spectrin, 150
Hydrophobic pocket, 249
Hydrophobic residues, in collagen, 148
α-Hydroxy acid oxidase, 12
Hydroxyacyl dehydratase, 377*t*
Hydroxyacyl-CoA dehydrogenase, 370
3-Hydroxyacyl-CoA dehydrogenase, 372*t*
Hydroxyapatite, 121
2-Hydroxybenzoic acid, 398*s*
3-Hydroxybutyrate dehydrogenase, β-hydroxybutyrate dehydrogenase, 373*t*
β-Hydroxybutyrate dehydrogenase, location, 13*t*
3-Hydroxybutyrate, β-hydroxybutyrate 37*t*, 370*s*, 371, 394, 396
22-Hydroxycholesterol, 393*d*
7-α-Hydroxycholesterol, 391*d*
Hydroxyisovalerate, in valinomycin, 181
Hydroxylysine, 505

3-Hydroxy-3-methylglutaryl-CoA, 388, 389*s*
Hydroxymethylglutaryl-CoA (HMG-CoA) synthase, 373*t*
Hydroxymethylglutaryl-CoA reductase, 388*d*, 389*d*
Hydroxyproline, 70*s*, 93*s*, 123*d*, 505
  in collagen, 93, 148
Hydroxyurea, 444*s*
Hyperammonemia, 455
Hyperchromic effect, in single-stranded DNA, 212
Hypoxanthine, 447*d*

Ideality, thermodynamic, 114
Idose, 28*s*
L-Iduronic acid, 52*s*
Imidazole, 75*s*
Immunodeficiency, 456
IMP, inosine 5′-monophosphate, 225*s*
Inbom errors of metabolism, 15
Induced fit hypothesis, 236, 273
Inducer, of gene, 508
Infant, 15
Inflammatory white cells, 1
Influenza virus, hemagglutinin, 112, 113*d*
Inherited disease, 21
Inhibition of enzyme:
  anticompetitive, 254
  competitive, 254
  mixed, 254
  noncompetitive, 254
  uncompetitive, 254
Inhibition pattern, 280
Inhibitor, 253
Initial reaction rate, 274
Initial velocity, 252
Initiation, of DNA replication, 460, 469
Initiation codon, 490*t*
  AUG, 510
Inner sheath, in cilia, 141, 142*d*
Inosine, 218*s*, 447*d*
Inosine 5′-monophosphate (IMP), 225*s*
Inositol, in phosphatidylinositol, 159t
Insulin, 394
  affect on glycolysis, 336
  gene, 518
  self association, 111, 112
  structure, 104, 506*d*
Integral protein, in membranes, 172
Integrin, 130
Interaction energy, 234
Intercalation into DNA, 473
Interconvertible enzyme, 301, 302
Intermediate density lipoproteins (IDL), 364
  definition, 168, 169*t*
Intermediate filament, 12, 131, 144, 145*t*, 151
Internal energy change (ΔU), 290, 292
International Union of Biochemistry and Molecular Biology (IUBMB), 241
International Union of Pure and Applied Chemistry (IUPAC), 241
Interphase, 140
Intervertebral disk, 122*t*, 127, 129
Intestinal epithelial cell, 363*d*
Intestine, 17
  secretory cell in, 17, 18*d*
Intracellular, 21
  compartmentation, 303
  symbionts, 12
Intracellular vesicle, size, 17
Intranucleolar chromatin, 14*d*
Intrinsic association binding constant, 271
Intrinsic binding constant, 287
  definition, 269
Intrinsic protein, in membranes, 172
Intron, 487, 497
Inuit, Eskimo, 419
Invertase, 52, 255
Iodate, 3
Iodine in, oligonucleotide synthesis, 211
Iodoacetate, modification of cysteine by, 69*s*
Ion-exchange chromatography, 77
  in amino acid analysis of proteins, 65, 66*d*
Ionic strength, 299, 305*d*
Ionizing radiation, 475
Ionophore, 181
Iron-sulfur protein/cluster:
  aconitase, 347
  in electron transfer chain, 405*t*
Iso mechanism, enzymatic reaction, 259
Isocitrate, 347*s*
  in the glyoxylase cycle, 356*s*
Isocitrate dehydrogenase, 13*t*
  activation by ADP and NAD^+, 351*d*
  inhibition by ATP and NADH, 351*d*
  reaction, 347*d*
Isocitrate lyase, 356*d*
Isoelectric point, glutamic acid, 64*t*
Isoenzyme, 321
Isolated (adiabatic) system, definition, 290
Isolation of genes, 476
Isoleucine, 283, 423*t*
  structure of, 54*t*
Isomerase, 230*t*
Isomerization, enzyme, 259
Isopentenyl pyrophosphate, 388*s*, 389*d*, 390*d*
Isoprene, 387
  acetate as precursor, 154
Isopycnic centrifugation, 5*d*
Isozyme, 285
IUBMB, International Union of Biochemistry and Molecular Biology, 241
IUPAC, International Union of Pure and Applied Chemistry, 241

Jansen, E. F., 239
Janus green B, 12
Joints, 17

K_a (acid dissociation constant), definition of, 58
K_a and K_b, relationship between, 59
Karyotype, 15
Katal, 257
K_b, basicity constant, 59
Keratan sulfate, 127, 128*d*
Keratin, attachment to fibroblasts, 145
Ketimine, 242*s*
Ketoacyl reductase, 377*t*
β-Ketoacyl synthase, 377*t*
Ketogenesis, 370*s*, 371, 373*d*
α-Ketoglutarate, 2-oxoglutarate, 420
Ketoheptose, 26
Ketone body, 370*s*, 371
  metabolism, 394*d*
  utilization in starvation, 394*d*
Ketone group, 42
Ketopentose, 26
Ketose, 25, 30
Ketotriose, 26
Kidney, 11*t*
Kinesin, 141*d*, 144
  in cellular motors, 143
Kinetic order, 274
Kinetic stability, 237
Kinetics of membrane transport, 175*d*
Kinetochore, 143*d*, 144
King, E. L., 258
King and Altman, procedure, 257, 258, 260*d*
Klenow fragment, of DNA polymerase I, 466
K_m, 253*d*, 256, 257, 261, 262, 262*d*, 264*d*, 268, 276–278, 281, 282, 284, 285, 287, 288, 289
  expression for, 260
KNF model, Koshland, Nemethy and Filmer, 273
Koshland, D. E., 273
Krabbe leukodystrophy, 383*d*
Krebs cycle, citric acid cycle, tricarboxylic acid, 345, 346
Krebs, H. A., urea cycle discovery, 435
K_w (ionic product of water) defined, 57
Kynurenine hydroxylase, 13*t*
  location, 13*t*

Lac operon, in *E. coli*, 516
α-Lactalbumin, in lactose synthesis, 115
Lactase, 330, 343
  conversion to pyruvate, 321*s*
  redox recycling, 322*d*
Lactate dehydrogenase, 249
  in Cori cycle, 327*d*
  estimation of $NADH/NAD^+$ ratio, 303*d*, 304
  isoenzymes of, 327
  in NAD^+ recycling, 321*d*
Lactate fermentation, 321
Lactate oxidase, 12
Lactobacillic acid, 182*s*
β-Lactoglobulin, 77, 78, 105
Lactonase, 339*d*
δ-Lactone, 41*s*
Lactose, 50*s*, 330
  synthesis, 115
Lactose (LAC) operon, 508*d*
  catabolic activator protein, 508
  in *E. coli*, 508
  lac operon, 508
Lactose synthetase, 115
Lagging, 466*d*
  strand, 457, 469*d*
Laminin, 131*t*
Lamins, 151
  in prophase, 145
Lanosterol, 389, 391*d*

LCAT, lecithin: cholesterol acyltransferase, 365*d*
LDL, low-density lipoprotein definition, 168, 169t
LDL receptor, 366*d*, 368, 390
Lead phosphate, 4
Lead zirconate titanate, piezoelectric effect of, 19
Leading strand, 467, 469*d*
Leeuwenhoek, Antonie van, 1
Lennard-Jones 6, 12 potential, 85
Lesch-Nyhan syndrome, 446
Leucine, 423*t*
  structure of, 54*t*
Leukemia, treatment, 444
Leukotriene:
  A_4, 387*d*
  B_4, 387*d*
  synthesis from arachidonic acid, 387*d*
Ligament, 121*t*, 125, 149
Ligand, 270
  binding, 115
Ligase, 230*t*, 241
Light micrograph, 4, 23
  of cells, 23
Lignoceric acid, 161
Limonene, 162*s*
  structure, 153*s*
Lineweaver-Burk plot, 253, 254*d*, 264*d*, 265, 266*d*, 277*d*, 278, 279, 285*t*, 286*d*
Link protein, in proteoglycan, 127, 128*d*
$\alpha(1 \rightarrow 4)$ linkage, 46
$\alpha(1 \rightarrow 3)$ linkage, 47
$\alpha(1 \rightarrow 6)$ linkage, 46
$\beta(1 \rightarrow 4)$ linkage, in cellulose, 47
Linoleic acid, 155*s*, 156, 265, 365, 377*d*, 395
  linoleate, 376, 377
α-Linolenic acid, 156*s*
Linoleoylcholesterol, 365*s*
Lipase, 11*t*, 362*d*, 363
Lipid:
  amphiphilic, 166
  classes of, 154*s*
  definition, 153*s*
Lipid bilayer, 7*d*, 167*d*
Lipid membrane, 12
Lipid-soluble vitamin, 363
Lipids of solvation, 183
Lipogenesis, 374, 394
Lipogenic tissue, 374
Lipoic acid, 352, 353*d*
  in pyruvate dehydrogenase, 116
Lipoprotein, 81
  HDL, 364
  IDL, 364
  LDL, 364
  overall structure, 169
  remnant, 365
  types, 169*t*
  VLDL, 364
Lipoprotein lipase, 365
5-Lipoxygenase, 387*d*
Liver, 9, 11*t*, 21, 23
  rat, 4
Liver cell, hepatocyte, 12, 17
Liver tissue, 17
Lock and key concept, Fischer's, 273
Locus control region, 509
Low-density lipoproteins (LDL), 364
  definition, 168, 169*t*
Lung, cilia in, 17
Ly system, of amino acid transport, 431
Lyase, 230*t*
Lysine, 233, 238, 242, 243*t*
  structure of, 56*t*
Lysolecithin, lysophosphatidylcholine, 160*s*
Lysophosphatidylcholine (lysolecithin), 160*s*, 365, 385*d*
Lysophosphoglyceride, 159
Lysosomal enzymes:
  cathepsin, 11*t*
  collagenase, 11*t*
  mammalian, 11*t*
  peptidase, 11*t*
  protease, 11*t*
  substrates, 11*t*
Lysosomal membranes, 4
Lysosome, 3, 6*d*, 9*d*, 11*t*, 12, 21
  primary, 22*d*
  secondary, 22*d*
  tertiary, 22*d*
Lysosomes in macrophages, 21, 22*d*
Lysozyme, 11*t*, 245
  active site, 246*d*, 247*d*
  molecular weight, 82*t*, 83*t*
  sedimentation coefficient, 82*t*
  substrate conformation, 245
  X-ray crystallography of, 245
Lysyl group, 2
Lysyl oxidase, 126
Lysylhydroxylase, 123
Lyxose, 28*s*

M phase of cell cycle, 461, 462*d*, 463
Macromolecule, energy for synthesis, 290
Macrophage, 10*d*, 11*t*, 21, 22*d*, 130
Major groove, DNA, 209*d*
Malaria, 148
Malate, 324*s*, 375*d*
  in the citric acid cycle, 348*s*
  in the glyoxylase cycle, 356*s*
Malate dehydrogenase, 13*t*, 375*d*
  cytoplasmic, 324
  mitochondrial, 348*d*
  (2) (oxaloacetate-decarboxylating)($NADP^+$), 375*d*
Malate isocitrate dehydrogenase, location, 13*t*
Malate synthase, 356*d*
Malate-aspartate shuttle, 333*d*
Malnutrition, 419, 457
Malonyl transacylase, 377*t*
Malonyl-CoA, 374, 376, 392, 394
Malonyl-phosphopantotheine, 376*d*
Malt extract, 228
Maltose, 45, 46, 51, 330
Mammalian, 11*t*
D-Mannitol, 41*s*
α-D-Mannopyranose, 38*s*
D-Mannose, 28*s*, 49, 51
  in gluconeogenesis, 330, 331*s*
Marfan syndrome, 126
Marker enzyme, 309
Mass action, principle of, 251, 269, 274
Mass-action ratio, definition, 299*d*
Mass spectrometry:
  determining protein molecular weight with, 84
  electrospray, 84
  matrix-assisted laser desorption (MALDI), 84
  plasma desorption, 84
Mast cell, 21
  staining, 21
Matrix of mitochondria, 11*d*, 12
Maximal velocity (V_{max}), 253
Melanocyte, 432
Membrane:
  asymmetric distribution of phospholipids, 173
  composition, 171*t*
  lipid, 12
  shuttle, 8
  structure, 171
  transition temperature, 173
  turnover, 21
Membrane flow model, 8
Membrane potential, 179
  definition, 407*d*
  measurement of, 407
Mental retardation, 15
Menten, M., 255
Mercaptoethanol, 7*s*, 68
6-Meracaptopurine, 444*s*, 445*d*
Messenger RNA (mRNA), 218, 489
Metabolic compartmentation, 300
Metabolic control theory, 310*d*
Metabolic pathway, control of, 265*d*
Metabolic shuttle, 333*d*, 334*d*, 415, 416
Metabolism, energy supply, 290
Metabolon, 442
Metachromatic leukodystrophy, 382*t*, 383*d*
Metaphase, 142
  chromosome, 144
Methanol, 232*s*
N^5, N^{10}-Methenyl tetrahydrofolate (5, 10, CH-THF) cyclohydrolase, 442
Methionine, 423*t*, 449*s*
  metabolism, 450*d*
  structure of, 55*t*
Methotrexate, 444*s*, 473
5-Methylcytosine, 199*s*
Methylene blue, 1, 2*ds*, 18
  in histology, 18
Methylene-THF reductase, 449
β-D-Methylfructofuranoside, 44*s*
β-Methylgalactoside, 51
α-D-Methylglucopyranoside, 44*s*
4-Methylpentanal, 393*d*
10-Methylstearic acid, 157*s*
N^5-Methyltetrahydrofolate, 449
Mevalonate kinase, 389*d*
Mevalonic acid, 389*d*
Micelle:
  in cholesterol synthesis, 165, 166*d*, 182*s*, 363*d*
  critical micellar concentration, 167
  ionic strength effect on, 167
  mixed type, 167
  size, 167
Michaelis constant, 253
Michaelis, L., 255, 256

Michaelis-Menten equation, 253, 254, 259
  data, 276*t*, 277*t*
Microbody, 7
Microcavitation, in sonication, 19
Microfilament, 131
Microorganism, 1
Microscope, 3, 255*t*
Microscope slide, 1, 255*t*
Microscopy, 17, 255*t*
  electron, 3, 255*t*
  light, 1
Microsome, 24
Microtubule, 12, 113*d*, 131, 141, 143, 150, 254, 255*t*
  ADP in, 141, 254, 255*t*
  associated protein (MAP), 140
  astral form, 144
  in cilia, 142*d*
  conformational change, 141
  dynamic instability, 143
  minus end, 141
  organizing center, 140, 254, 255*t*
  organizing complex (MTOC), 143, 254, 255*t*
  plus end, 141, 143, 144
  treadmilling, 143
Microvillus, microvilli, 17, 136*d*
Miescher, F., 198
Mineralocorticoid, 393*d*
Minor groove, DNA, 209*d*
Minus end, of tubulin, 140
Mitchell, P., 409, 410
Mitochondria, 5, 9, 12, 13*t*, 16, 17, 21
  in adipocyte, 16*d*
  ATP translocase in, 241
  cristae in, 11*d*, 12
  DNA in, 11*d*, 12
  enzyme distribution in, 13*t*
  inheritance of DNA in, 24
  inner membrane, 13*t*
  in macrophages, 21, 22*d*
  matrix, 11*d*, 12
  nucleic acid in, 12
  outer membrane, 13*t*
  reproduction, 110
  RNA polymerase of, 494
  RNA translation, 507
  size, 17
  structure, 11*d*, 12
  supravital stain of, 12
  transfer of acetyl groups from, 375*d*
Mitochondrial DNA, 24, 461
Mitochondrial matrix, 13, 19
  staining, 19
Mitochondrial membrane, 8
Mitochondrial nucleic acids, 12
Mitochondrion, 6*d*, 16*d*, 20*d*
  being phagocytosed, 20*d*
Mitosis, 150
Mitotic spindle, 12, 108, 142, 143*d*, 151
Mitrotubule, 139
Mixed inhibition, 254, 255*t*, 264, 280*d*
  of enzyme, 263*d*
Mixed micelle, 363*d*
Mixed-function oxidase, 426
Mobile carrier mechanism of membrane transport, 180*d*, 181
Models of binding cooperativity, 267
Mole fraction, 307
Molecular distortion, 233
Molecular mass:
  definition of, 81
  of proteins, 81
  rule of thumb for proteins, 81
Molecular orbital calculations, 244
Molecular pump, 7
Molecular strain, 233, 234, 235, 247
Molecular weight:
  definition of, 81
  of proteins, 81
Molecularity, 274
  definition, 252
Mongolism, 15
Monoacylglycerol, 362*s*, 363*d*, 366*s*
2-Monoacylglycerol, 362*s*, 378*d*
Monoamine oxidase, location, 13*t*
Monod, J., 270
Monod, Wyman and Changeux (MWC) model, 271
Monolayer of lipids, 165, 166*d*
Monooxygenase, 390
  NADPH cosubstrate, 392
Monosaccharide, 25, 37*s*
Monosialoganglioside, 383
Motor protein, 140
mRNA, messenger RNA, 489
  capping, 497
  consensus sequence in, 497
  exon, 497
  intron, 497
  leader region, 503
  polyadenylation, 497
  polycistronic, 512
  processing, 497, 498*d*
  processing enzymes, 497
  Shine-Dalgarno sequence, 503
MTOC, microtubule organizing complex, 142*d*, 143
Mucopolysaccharidase, 11*t*
Mucopolysaccharide, 11*t*, 47
Mucus-secreting cell, 17, 18*d*
Multienzyme complex, 116
  assembly, 145
  metabolic advantages, 115
Multiforked chromosome, 263*d*, 482, 483
Multifunctional enzyme, (DNA polymerase I), 465
Multinucleated cells, 17
Multireactant enzyme, 259
Multisubstrate enzyme, 259
Muscle, 108, 138, 139, 144, 145
  myofibril, 138*d*
  myosin, 137
  rigor, 138
  skeletal, 17
  tropomyosin, 138
  troponin, 138
Muscle cell, myocyte, 17
Muscular dystrophy, 131
Muscularis mucosae, 427*d*
Mutarotation, 34, 49*d*
Mutation, in DNA, 458
MWC equation, model, 173*d*, 270, 287
Myocyte, muscle cell, 17
Myofibril, 137
  muscle, 138*d*
Myofilament, sliding filament model, 137
Myoglobin, 102, 118, 267, 452
  globin fold, 118
  hemoglobin comparison, 151
  hydrophobic interactions, 118
  molecular weight, 78
  specific optical rotation, 104
  x-ray analysis, 97
Myosin, 136*d*, 137, 139, 150
  α-helical tail, 137
  actin binding to, 138*d*
  actin interaction, 133
  ATP and, 133
  binding to actin, 135
  coiled-coil, 137
  gel-to-sol transition, 135
  gelsolin and, 135
  light chain kinase, 138
  muscle, 137
  phosphorylation, 138
  sarcomere, 138*d*
  thick filament, 137, 138*d*
Myristic acid, 185
Myristoleic acid, 185
Myristoyl-CoA, 371*d*

Na^+ symport, 430
Na^+/H^+ antiport, in mitochondria, 409
Na^+/K^+ ATPase , 430
NAD^+:
  1, 3-bisphosphoglycerate, 240
  in pyruvate dehydrogenase, 116
  structure, 311*d*
NAD^+/NADH:
  conjugate redox pair, 403
  hydride ion in, 403
  light absorbance at 340 nm, 403
NADH, 240
  in fatty acid oxidation, 370
  structure, 311*d*
NADH/CoQ oxidoreductase, 405*t*
$NADP^+$, NADPH:
  in the pentose phosphate pathway, 339*d*, 340
  structure, 312*s*
NADPH, in fatty acid synthesis, 376, 377*t*
Nail, finger, 145
Nalidix acid, 474, 475*s*
Nasal lining, cilia in, 17
Nascent fragments, 466
Nebulin, and vernier principle, 138
Negative control of a gene, 508
Negative cooperativity, 269, 284, 285
  enzyme kinetics and binding, 266
Negative effector, of aspartate carbamoyltransferase, 229
Negative feedback inhibition:
  in the citric acid cycle, succinate dehydrogenase, 351
  of enzyme, 265*d*, 283*d*
Negative feedforward control, 284
Negative nitrogen balance, 419
Negative supercoiling of DNA, 468
Nemethy, 273
Nernst potential equation, 307
Nernst redox equation, 297*d*, 406*d*
Nerve gas, diisopropyl fluorophosphate, 239*s*

Nerve, squid, 1
Nervonic acid, 161
Neurofilaments, 145*t*
Neurospora, cytochrome *c*, 98*td*
Neurotransmitter catabolism, location, 13*t*
Nexin, in cilia, 141, 142*d*
Nick in DNA, 464
Nick translation of DNA, 467
Nickel, in urease, 241
Nicotinamide adenine dinucleotide, (NAD), 311*s*, 312
Niemann-Pick disease, 382*t*, 383*d*
Ninhydrin, reaction with amino acids, 65*s*
Nitramide, 233*s*
Nitrite, 418
Nitrogen:
  atmospheric, 419
  balance, 419
  fixation, 419
  fixing bacteria, 420
Nitrogenase, 419*d*
  ATP in, 420*d*
  flavoprotein in, 420
  $\Delta G°$, 419
  H_2 in, 420*d*
NMR (nuclear magnetic resonance spectroscopy), 100
  of protein, 239
Nobel prize 1975, 13
Nojirimycin, 52
Noncompetitive inhibition, 254, 255*t*, 264, 280*d*, 289
  of enzyme, 263*d*
Nonequilibrium reaction, 299
Nonessential amino acid, 423*t*
Nonhomologous recombination of DNA, 463
Nonideality, 305
  thermodynamic, 114
Noninteracting binding sites, 269
Nonlinear least squares regression, 254
Nonprotein amino acid, 56
Nonsense codons, 491
Norepinephrine, 432, 433*s*
Novobiocin, 474
Nuclear envelope, 7, 14*d*, 142, 151
Nuclear lamin, 145*t*
Nuclear lamina, composed of lamin, 145
Nuclear magnetic resonance spectroscopy, NMR, 100
Nuclear magnetic resonance spectroscopy (NMF), of protein, 239
Nuclear membrane, 6*d*, 14*d*
Nuclear pore, 14*d*
Nuclease, 11*t*, 219
Nucleic acid, 198
  breakdown, 447*d*
  degradation, 446, 447*d*
Nuclein, 198
Nucleoid, in bacteria, 489
Nucleolus, 6*d*, 14*d*
Nucleophile, 232*d*
Nucleophilic attack, 231, 232*d*, 249
  in DNA synthesis, 465
Nucleoside, definition, 201
Nucleoside, diphosphokinase, 13*t*, 348*d*
  location, 13*t*
Nucleosome, 217*d*, 473
Nucleotidases, 446
Nucleotide, 202
Nucleus, 1, 2*d*, 14*d*, 16*d*, 17, 343
  of actin, 132*d*
  in adipocyte, 16*d*
Numerical methods, in kinetic analysis, 254

O_2, transport in erythrocytes, 15
Obligatory order, enzymatic reaction, 259
Octadecanoyl-CoA, 369*s*
Ogston, A. G., 359
Ogston's three-point attachment proposal, 359*d*
Oiled coil, elastin, 126
Okasaki fragments, 466, 467, 470
Oleate, 369*s*, 376
Oleic acid, 156*s*
  oleate, 369*s*, 376
Oleoyl-CoA, 369*ds*
Oligomer, 270*s*
Oligomeric protein, 273
Oligomycin-sensitivity-conferring protein (OSCP), 412
Oligonucleotide, 11*t*, 25, 206*s*, 247
  definition, 44
  staining with PAS, 21
Oligonucleotide synthesis, solid phase, 211*d*
OMP decarboxylase, 438
Open system, definition, 290
Operator, of gene, 508
Opposite polarities of DNA strands, 466
Optical activity, 52
  definition of, 94
  of proteins, 94, 95
  sugars, 26, 27
Optical rotation, 44, 52
Optical rotatory dispersion:
  definition of, 94
  of proteins, 94, 95*d*
Optimum pH, of enzyme, 261, 262*d*
Order of reaction, 274
  definition, 252
Ordered Bi Bi:
  enzymatic reaction, 260*d*
  reaction, 280
Organic sulfate, 11*t*
OriC, 318*d*, 460, 463, 470
Origin of the chromosome (oriC), 460
Origins of replication, 472
Ornithine, 70*s*, 230*s*
L-Ornithine, 435*d*
Ornithine carbamoyltransferase, 435
  location, 13*t*
Orotate, 437
Orotic aciduria, 456
Osmium tetroxide, 3
Osmotic lysis, of cells, 19
Osmotic pressure, 81, 309*d*
Ostwald's dilution principle, 147
Outer membrane, 13*t*
  mitochondrial, 13*t*
Ovalbumin, 102
  molecular weight, 83*t*
Overall reaction order, 252
Ovomucoid, molecular weight, 83*t*
Ovum, 1
Oxaloacetate, 324*s*, 325*s*, 374*s*, 375*s*, 422*s*
  in the citric acid cycle, 346*s*, 350*s*
  in gluconeogenesis, 323*s*
  in pyruvate carboxylase reaction, 353*s*
Oxalosuccinate, 347*s*
Oxidase, mixed-function, 426
Oxidation of fatty acids, 368
Oxidation reaction, 295
Oxidation-reduction, 230*t*
Oxidative deamination, 124
Oxidative phosphorylation:
  chemiosmotic model, Mitchell's, 408
  definition, 402
  mechanistic models of, 408
Oxidoreductase, 230*t*, 240
2,3-Oxidosqualene: lanosterol synthase, 391*d*
2-Oxoacid, 423*t*
2-Oxoacid dehydrogenase (BCOADH), 432
  location, 13*t*
3-Oxoacid transferase, 374
2-Oxoglutarate, 240*s*, 422*s*
  in the citric acid cycle, 347*s*
  α-ketoglutarate, 420
  product of prolylhydroxylase, 123*d*
2-Oxoglutarate dehydrogenase complex:
  product inhibition by NADH and succinyl-CoA, 351
  reaction, 347*s*
Oxolinic acid, 475*s*
Oxonium ion, 233*s*, 233*d*
6-Oxy-2-aminopurine, 200
Oxygen, 12, 15
  in oxygenase reactions, 12
Oxygen 18, label, 413
Oxygen affinity, hemoglobin, 147, 148
Oxyhemoglobin, 118

Palade, G., 4, 13
Palade's granules, 13
Palmitate, 369*s*
  structure, 153*s*
Palmitic acid, 155*s*, 156, 376, 396
  palmitate, 376
  synthesis of, 374
Palmitic acid synthesis, 376, 377*t*
Palmitoyl thioesterase, 377*t*
Palmitoyl-CoA, 369*ds*, 370*d*, 371*d*, 372*t*, 395
Palmitoylcarnitine, 370*d*
Pancreas, 506
Pancreatic lipase, 362, 363*d*
Pancreatic peptidases, 248
Pantothenic acid, 56
  in CoA, 345*s*, 346
Paracrystalline structures, 18
Parental DNA, 459
Parietal cell, 427*d*
Partial reaction, linked to ATP, 299*d*
Partial specific volume, 81
Partitioning, of chromosome, 458
PAS staining, 21
  of erythrocytes, 21
Pasteur effect, 335*d*, 418
  connection to oxidative phosphorylation, 418

Pasteur effect (*Cont.*):
effect on oxidative phosphorylation, 412
Pasteur, L., 335
Payen, 228
PCR, polymerase chain reaction, 477–479, 480, 481*d*, 487
Peas, 419
Pectin, 47
Pentamicrofibril, 124*d*
tropocollagen, 123
Pentose, 25
Pentose phosphate pathway, 16, 241
definition, 339
in erythrocytes, 16
Pepsin, 229, 248, 426
Peptic cell, 427*d*
Peptidase, 11*t*
Peptide, 53–75
dipeptide, 67
enzymatic cleavage of, 80
glutathione, 67*d*
naming of, 67
oligopeptide, 67
polypeptide, 67
N-terminal, 67
tripeptide, 67
Peptide bond, 66, 67, 231*td*, 427
structure of, 67*d*
Peptide hormones, 427
Peptidyltransferase, 504*d*
Perhydrocyclopentanophenanthrene, 163*s*
Periodic acid Schiff (PAS), 2*d*
staining, 21
Peripheral protein, in membranes, 172
Permease, 174
Peroxidase, 385*d*, 452
Peroxisome, 12
Persoz, 228
pH, definition of, 57, 58
pH effect, on enzyme rate, 261*d*, 262*d*
Phagocytic vacuole, 10*d*, 22*d*
in macrophage, 22*d*
Phagocytosis, 21
Phagosome, 10*d*, 19
Phenanthrene, 163*s*
Phenol, 249
Phenomenase, 241
Phenylacetate, 455*s*
Phenylalanine, 244, 423*t*, 425*s*
hydroxylase, 425*d*
structure of, 55*t*
Phenylisothiocyanate:
Edman reagent, 79*s*
in protein sequencing, 79*d*
Phenylketonuria, 426, 455
Phenyllactate, 455*s*
Phenylpyruvate, 426, 455*s*
ϕX174:
genome size, 215, 216*t*
single-stranded circular DNA, 226
viral DNA, 510
Phosphatase, 11*t*
Phosphate, 3*sd*, 19
ester, 2, 233
esters of sugars, 42
product of phosphatase reaction, 19*d*
Phosphatidic acid:
phosphatidate, 378
synthesis of, 380*d*
Phosphatidylcholine: cholesterol acyltransferase (LCAT), 159*t*, 365, 379, 380*d*, 381*d*, 382*d*, 385*d*, 397
Phosphatidylethanolamine, 159*t*, 381*d*
Phosphatidylglycerol, 159*t*
Phosphatidylinositol, 159*t*
Phosphatidylserine, 159*t*, 381*d*
Phosphocreatine, 456*s*
Phosphodiester linkage, 204*s*
Phosphodiesterase, 338, 446
Phosphoenolpyruvate:
carboxykinase (PEPCK), 324, 325*d*
in gluconeogenesis, 325*d*
in glycolysis, 316
Phosphoenzyme, 232*d*
Phosphoethanolamine, 381*d*
Phosphofructokinase, 283, 360
ADP: positive heterotropic effector, 284
ATP: negative homotropic effector, 284
citrate: negative heterotropic effector, 284
control of, 301*d*
control in gluconeogenesis, 326*d*
control point in glycolysis, 318, 319*d*
in glycolysis, 313
thermodynamics of, 300*td*
Phosphoglucomutase, 327, 329
in glycogen synthesis, 327, 360
thermodynamics of, 294
6-Phosphogluconate, in the pentose phosphate pathway, 339*s*
Phosphogluconate pathway, 339
6-Phosphogluconolactone, in the pentose phosphate pathway, 339*s*
2-Phosphoglycerate, in glycolysis, 315*s*
3-Phosphoglycerate, 425*d*
in glycolysis, 315*s*
in substrate level phosphorylation, 402*s*
Phosphoglycerate dehydrogenase, 425*d*
Phosphoglycerate kinase, 315*d*
Phosphoglycerides:
major classes, 159*t*
structure, 158*s*
2-Phosphoglycerol, 3*d*, 4, 11*t*
Phosphoglyceromutase, 315*d*
3-Phosphohydroxypyruvate, 425*d*
Phospholipase, 362, 363
Phospholipase A_1, 363, 381
Phospholipase A_2, 363*d*, 381, 385*d*
Phospholipase C, 381
Phospholipase D, 381
Phospholipid, 11*t*
absorption, 362
definition, 158
synthesis, 379
Phosphomannose isomerase, 330, 331*d*
Phosphomevalonate, 389*d*
kinase, 389*d*
Phosphomonoester, 11*t*
Phosphopantetheine, 376
Phosphophosphorylase kinase, 338*d*
Phosphoprotein phosphatase, in pyruvate dehydrogenase control, 352, 354*d*
5-Phosphoribosyl, 1-pyrophosphate (P-Rib-PP), 437, 438, 440*d*, 440*s*
Phosphoribosyltransferases, 446
Phosphoryl-group transfer, thermodynamics of, 298*d*
Phosphorylase, *a*, *b*, 115
Phosphorylase *a*, from '*b*' form, 337, 338*d*
Phosphorylase *b*:
activation to '*a*' form, 337, 338
thermodynamics of, 295
Phosphorylase kinase, 115, 337*d*
Phosphorylase phosphatase, 115
Phosphorylation, 145
in pyruvate dehydrogenase, 116
Phosphorylserine, 239
3-Phosphoserine, 425*d*
Phosphoserine phosphatase, 425*d*
Phosphoserine transaminase, 425*d*
Phosphosphingolipid, 161*s*
Phosphotungstic acid, staining of tropocollagen, 124*d*
Photoreactivation of DNA, 476
Phylogenetic scheme, 14
Piezoelectric effect, 19
Ping pong, enzymatic reaction, 259
Pitch of helix, 107
Placenta, oxygen transfer across, 147
Planck, M., 238
Plant, 1
Plasma membrane, 6*d*, 7*d*, 9*d*, 10*d*, 16, 19, 21, 129, 134, 150
in adipocyte, 16*d*
cilia, 141, 142*d*
Plasmalogen, 159
Plasmid, 217, 461
Plasmodium, 1
Plus end, of tubulin, 140
Pointed ends, of actin filaments, 139
'1080' poison, fluoroacetate, 359, 360*s*
Polar, helical rods, 131
Polarimeter, 26, 52
Polarity, actin, 133
Polio virus, 110
Poly(A) polymerase, 497
Poly(A) tail on RNA, 497
Poly(α-D-glucosamine), chitosan, 51
Poly-L-glutamate, 103
Polyadenylation of RNA, 497
Polyaminoacids, 107
Polycistronic mRNA, 494, 512
Polydeoxynucleotide, 206*s*
Polyhydroxy-aldehyde, 25
Polyhydroxy-ketone, 25
Polyisoprenoid, 163
Polyketides, acetate as precursor, 154
Polymer, of glucose, 47
Polymerase chain reaction (PCR), 477–481*d*, 487
Polynucleotide, 204*s*
Polyproline, 93
Polyproline helix, 93
Polyribosome, polysome, 505
Polysaccharidase, 11*t*, 25, 44, 46, 126
as a substrate, 11*t*
Polysome, polyribosome, 13, 505
Polyunsaturated fatty acid, 155*s*
Pore, 180*d*
Porin, 174

Porphobilinogen, 451*s*
  synthase, 451
Porphyria, 457
  synthesis, 451
Positive control of a gene, 508
Positive cooperativity, 269, 273*d*, 284
  enzyme kinetics and binding, 266
Positive feedback control, 284
Positive supercoiling of DNA, 468
Positively cooperative binding, 286
Posttranslational modification, of proteins, 505, 515, 516
Potentia hydrogenii, pH, 57
Potter-Elvehjem homogenizer, 5
Pregnenolone, 392*s*, 393*d*
Prephenate, 244*s*
Preproinsulin, structure, 506*d*
Pribnow box, 511, 512, 517
  DNA, 492*d*
Primary active transport, 179*d*
Primary lysosome, 10*d*, 20*d*, 22*d*
  in macrophage, 22*d*
Primary transcript, 493
Primase (DnaG), 467, 469*d*
Primer, DNA, 465, 479
Primidine nucleotide, binding to aspartate transcarbamoylase, 115
Primosome, 467
Principal of mass action, 251, 274
Procarboxypeptidase, 427, 428*d*
Prochirality, 359
Procollagen:
  assembly, 149
  peptidases, 123
Proelastase, 248*d*, 427
Progestogen, 393*d*
Progress curve, 252
Proinsulin, structure, 506*d*
Proline, 423*t*
  L-, D-, 236*s*
  structure of, 55*t*
Proline racemase, 236
  planar and tetrahedral intermediates, 236
  proline as substrate, 236*s*
Prolylhydroxylase, 505
Promoter, 140, 487, 492
  binding, 115
  of gene, 508
  mobile, 467
  self association, 112*d*
Proofreading of DNA, 469
Prophase, lamins in, 145
Propinquity, effect in enzymes, 230, 242
Propionate, structure, 153*s*
Prostacyclin, 384*d*
Prostaglandin, 376, 398
  nomenclature, 383
  PGD_2, 384*s*, 386*d*
  PGE_2, 384*s*, 386*d*
  PGF_2, 153*s*, 384*s*, 386*d*
  PGG_2, 385*ds*, 386*d*
  PGH_2, 385*d*, 386*ds*
  PGI_2, 384*s*, 386*d*
  structure, 383
  structures, 384*d*
  synthase, 385
Prostanoic acid, 383*d*
Prostate glands, 4
Prosthetic group, 81, 232
Protease, specificity, 11*t*, 426, 429*t*
Protein, 2, 76–107, 108, 117, 228, 386*d*
  α-helix, 112
  affinity chromatography of, 77
  allosteric interactions, 118
  conformational change, 115
  conformational changes in, 109
  conformational entropy, 84
  connective tissues, 108
  cytoskeletal, 108
  denaturation with urea, 81
  dephosphorylation, 115
  dephosphorylation, in pyruvate dehydrogenase, 116
  dialysis of, 77
  disulfide bonding in, 109
  Edman degradation of, 78, 79
  electrophoresis of, 77
  electrostatic interactions, 84, 85*d*, 86*t*
  folding, 84, 108
  functions of, 76
  hydrogen bonds, 85, 86*t*
  hydrogen bonds in, 108
  hydrophobic interactions, 85, 86*t*, 108
  ionic bonds in, 108
  packing contacts, 117
  phenylthiohydantoin derivative, 79*s*
  phosphorylation, 115
  phosphorylation, in pyruvate dehydrogenase, 116
  protein folding, 108
  purification of, 76–78
  quaternary structure, 81, 108, 117
  salt bridge, 85*d*
  salting out of, 77
  sequence homology, 97, 98
  sequencing of, 78–81
  sliding contacts, 117*d*
  solubility of, 77
  structure dictated by sequence, 87
  supramolecular structure, 108
  van der Waals interactions, 85, 86*t*
Protein 4.1, 136, 137*d*
Protein channel, 7*d*
Protein conformation, of enzyme, 239
Protein degradation, 430*d*
Protein denaturants, 87
Protein disulfide-isomerase, 241
Protein evolution, phylogenetic tree, 97, 98*d*
Protein kinase (A), 337*d*
Protein structure, 87–89
  accessing via world wide web, 99
  β-α-β motif, 97*d*
  determination of, 99, 100
  domain structure, 96
  families of, 96
  folding domains, 97
  hierarchy of, 95
  planarity of peptide bond, 88
  primary, 95
  quaternary, 96, 97
  Ramachandran plot, 89*d*
  secondary, 95
  stylized representation of, 94*d*
  supersecondary, 95
  tertiary, 95
Protein structure (*Cont.*):
  topology diagrams, 96*d*
Protein synthesis, 13
  elongation, 500, 504*d*
  enzymes, 241
  inhibitors, 500
  initiation, 500, 504*d*
  release factors, RF1, RF2 and RF3, 505
Protein turnover, 431
Proteoglycan, 121, 126, 127, 128*d*, 129
  aggregated, 129*d*
Proteoliposome, 170, 171
Proteolysis, 145
  in virus assembly, 110
Proteolytic enzyme, 148, 229, 426
Protofilament, of tubulin, 139
Protomer, 111
  of binding protein, 270
Proton motive force:
  $\Delta\mu_{H^+}$, 407
  definition, 407*d*
Proton translocation in mitochondria:
  by complex I, 410
  by complex III, 410
  by complex IV, 410
  loop mechanism, 410*d*
  not by complex II, 410
  proton-motive Q cycle, 410, 411*d*
Proton-motive Q cycle, 410
Protoporphyrin IX, 451, 452*s*
Proximal histidine, in hemoglobin, 147
Proximity, effect in enzymes, 230
Pseudo substrate, 236
Pseudogene, 499
Pseudomonas, bongkrekic acid from, 415
Pseudouridine, 218*s*
Purine, 198*s*
  de novo synthesis, 440*d*
  inhibition of de novo synthesis, 445*d*
  metabolism, 437
  salvage synthesis, 446*d*
Purine tautomer:
  enol, 200*s*
  imino, 200
  keto, 200*s*
Puromycin, 507*s*, 516
Pyran, 35*s*
Pyranose, 49
Pyranoside, 44*s*
Pyridine, 75*s*
Pyridoxal phosphate, 232, 421*s*, 454
Pyridoxamine phosphate, 421*s*
Pyrimidine, 198*s*
  bifunctional enzyme, 438
  biosynthesis by *E. coli*, 229
  CAD protein, 438
  de novo synthesis, 438, 439*d*
  metabolism, 437
  pathway in *E. coli*, 438
  salvage synthesis, 446*d*
  synthesis: control of, 266
  trifunctional protein, 438
Pyrophosphatase, 465
  in glycogen synthesis, 328
Pyrophosphate, 465
Pyrophosphomevalonate decarboxylase, 389*d*
Pyrrole 2-carboxylate, 236*s*

Pyruvate, 240*s*, 375*d*, 424*t*
complete oxidation of, thermodynamics, 320*d*
conversion to acetyl CoA, 346
conversion to ethanol, 322*sd*
conversion to lactate, 321*s*, 322*d*
in glycolysis, 311
inhibition of membrane transport of, 184
metabolic fate of, 319, 320*d*
in pyruvate carboxylase reaction, 353*s*
Pyruvate carboxylase, 323*d*
allosteric activation by acetyl-CoA, 353, 354*d*
reaction, 353
subunit association, 309
Pyruvate decarboxylase, 116, 117, 322, 394, 454
acetyl-CoA as substrate, 116
CoA as substrate, 116
(E_1) reaction, 352, 353*d*, 354*d*
NAD^+ as substrate, 116
pyruvate as substrate, 116
subunit complex, 116, 117*d*
Pyruvate dehydrogenase complex, 352–354
control of, 302*d*, 352
inactivation by phosphorylation, 352, 354*d*
mitochondrial location, 352
overall reaction, 352*s*
Pyruvate dehydrogenase kinase, 352, 354*d*
Pyruvate kinase:
activation, 318
control point in glycolysis, 318
in glycolysis, 316*d*
inhibition, 318
positive feed forward control of, 318

Δq, heat change, 292
Q-structure, 482
Quad, reaction, 259
Quaternary structure, 111, 117
hemoglobin, 147

Rabbit, cytochrome *c*, 98*td*
Racemase, active site, 236
Racemic mixture, 26, 241
Radial spoke, in cilia, 141, 142*d*
Radioactive labels, study of citric acid cycle, 358
Ramachandran plot, 89*d*
Rapid-equilibrium reaction, 281
Rapoport-Luebering shunt, 344
Rat liver, 4
Rate constant, 237
definition, 251
Rate enhancement:
in enzymatic reaction, 233
enzyme, 237, 238, 241, 249
Rate of reaction, 251
Rate-determining step, 299
Reactancy, definition, 259
Reaction coordinate, 235*d*
Reaction order, 274
data, 275*t*
plots, 275*d*
Reaction rate, 251
Receptor, 129
Recombinant DNA, 476, 477, 478, 487
technology, 476
Recombinase, 471
Rectangular hyperbolic, 252, 253
Red blood cell, 15, 19, 21, 22, 23, 108, 148, 228, 289
cytoskeleton, 136, 137*d*
erythrocyte, 15, 16*d*, 119
hemoglobin in, 267
human, 22
life span, 21
lipid bilayer experiment, 185
number of hemoglobin molecules in, 22
PAS staining, 21
staining, 19
volume, 22
Redox potential, 403
Redox potentiometry, 406, 407*d*
Redox reaction, 295
Reducing equivalents, 313
Reduction potential, 297
Reduction reaction, 295
Regeneration, of cytoplasmic NAD^+, 332
Regression analysis, initial estimates, 254
Regulation:
of the citric acid cycle, 350*d*, 351*d*
of lipid metabolism, 392
Regulation of cholesterol synthesis, 390
Regulatory enzyme, 266*d*, 283, 286, 287
Regulatory subunits, of aspartate carbamoyltransferase, 229
Reinitiation of replication, 482
Remnant, lipoprotein, 365
Repair of DNA, 458, 475
Repair synthesis of DNA, 458, 486
Replication, 458
of DNA, 458
initiation of, in bacteria, 469
Replication bubble, 460
Replication of DNA:
bidirectional, 460
bubble, 460*d*
control of, 462, 482
D loop in, 461
discontinuous, 467*d*, 484
elongation, 472, 483, 486
enzymology of, 464
fork(s), 459, 460*d*
helix-destabilizing protein in, 468
inhibitors of, 473
initiation, 460, 464, 472, 483, 486
lagging fragments in, 466*d*, 467
nascent fragments in, 466*d*, 467
nick sealing, 485
Okazaki fragments in, 466*d*, 467
origin of, 460, 464
rate of, 463
reinitiation of, 463, 482
rolling-circle mechanism, 463, 464*d*
stages of, 461*d*
termination of, 460, 464, 483, 486
termination in eukaryotes, 472*d*
Replication fork, DNA, 458, 459, 460*d* 463*d*
Replication fork trap, 471*d*
definition, 485
Replicon, 460, 461, 462*d*, 463, 464, 472
Repressor protein, 508
Reptile, 455
Residual body, 20*d*
of lysosome, 20*d*
Residue, in sugars, 44
Resolution, 3
Respiration, 349
Respiratory chain:
electron transport chain, 13*t*, 402
location, 13*t*
Respiratory tract, 17, 417, 418
cilia in, 17
mucous secretion in, 17
Restriction endonuclease, 219, 477
Reticulin, 122*t*
Reticulocyte, 19, 23, 24, 509, 516
rate of hemoglobin synthesis in, 23
staining, 19
Retro-aldol condensation, 314
Reverse transcription 479
Reverse turn, in elastin, 125
Rhamnose, 51
L-Rhamnose, 40*s*
Rhizobium, 420
Rhodanese, 229
Rhodopseudomonas spheroides, 283
D-Ribitol, 49*s*
L-Ribitol 1-phosphate 49, 50*s*
D-Ribitol 5-phosphate, 49, 50*s*
β-D-Ribofuranose, 38*s*
Ribonuclease (RNase), 219, 446, 447*d*
active site, 229
substrate specificity, 229
Ribonucleic acid, RNA, 198
Ribonucleoside, 201, 202*s*
Ribonucleotide, 202
reductase, 442*d*
β-D-Ribopyranose, 39*s*
β-D-Ribose, 28*s*, 49, 51, 198
D-Ribose 5-phosphate, in the pentose phosphate pathway, 340*s*
Riboside, 43
Ribosomal RNA, rRNA, 14, 109, 218, 491
genes, 463
Ribosome, 6*d*, 7*d*, 8*d*, 13, 17, 19, 502
bacterial, 502*t*
eukaryotic, 502*t*
genes for, 498
size, 17
staining, 19
subunit structure, 502*t*
Ribosome subunits, 109*d*
in *E. coli*, 109*d*
sedimentation coefficients, 109*d*
D-Ribulose, 30
D-Ribulose 5-phosphate, 396
in the pentose phosphate pathway, 340*s*
L-Ribulose 5-phosphate, 240*d*
Ribulose phosphate epimerase, 241
Ricinoleic acid, 185
Rifampicin, 470, 500, 517
RNA, ribonucleic acid, 11*t*, 13, 19, 114, 198, 218
mRNA, 218
rRNA, 218
staining, 2, 19
structure, 218

RNA, ribonucleic acid (*Cont.*):
as a substrate, 11*t*
tRNA, 218
5S RNA, 494
RNA polymerase, 467, 470, 485, 491*d*, 512, 513
I, 494
II, 494, 495
III, 494
bacterial, 494, 500
mitochondrial, 494
subunit structure, 492, 493*d*
RNA primers, 467
RNA transcription:
complex, 495*d*
factors, 495*t*
TFIIA, TFIIB, TFIID, TFIIE, 495*d*
RNA translation, 514, 515
bacterial, 514
leader region, 514
RNase, ribonuclease, 219
Rolling-circle model of DNA replication, 463, 464
Romanowsky dye, 1
Rotenone, inhibition of electron transport chain, 406*d*
Rotenone-insensitive NADH, cytochrome *c* reductase, 13*t*
Rough endoplasmic reticulum (RER), 4, 8*d*
Royal Society of London, 1
RRNA, ribosomal RNA, 491
Rubber, self lubrication, 126

S, entropy, 290
ΔS, entropy change, 292
S phase of cell cycle, 461, 462*d*, 463, 473
S—S bond, 240*s*
Saccharomyces cerevisiae, 472
Sacromere, 139
Salt bridge:
in electrochemical cell, 296, 297*d*
in proteins, 85*d*
Salt links, in hemoglobin, 121
Salvage synthesis of nucleotides, 445
Sandhoff disease, 383*d*
Sanger, F., 473
Sarcomere, 137
Sarcoplasmic reticulum, 17
Satellite, DNA, 15
Saturated fat, 419
Saturated fatty acid, 155*s*
Schiff base, 3*s*, 125*d*, 242, 454*d*
Second law, of thermodynamics, 292*d*
Second messenger, cAMP, 337
Second-order reaction, 274
Secondary active transport, 179*d*
Secondary amino acids, proline, hydroxyproline, 93
Secondary lysosome, 10*d*, 22*d*
Secondary plot, noncompetitive inhibition, 279*d*
Secondary redox mediators:
ascorbate, 406
definition, 406
Secretin, 427
Secretion, of chylomicrons and VLDL, 379*d*
Secretory:
cells, 17
granules, 506
protein, 4
vacuole, 4, 8, 9*d*
Sedimentation analysis, 81
Sedimentation coefficient, 81
Sedoheptulose 7-phosphate, 341*s*
Segregation, of chromosome, 458
Self-assembly of viruses, 110
Self-associating proteins, 147
Self association:
of protein, 111, 112
spectrin, 150
Semiconservative replication, of DNA, 459, 481
Semiquinone, 404
Sense strand, coding strand, 489
Sephadex, 78
Sequence, 458
Sequential, enzymatic reaction, 259
Sequential feedback:
control of metabolism, 283*d*
on enzyme, 283*d*
Sequential ordered, enzymatic reaction, 259
Sequential random, enzymatic reaction, 259
Serine, 423*t*, 424*s*
hydroxymethyl transferase, 424, 449
phosphate, 74*s*
in phosphatidylserine, 159*t*
in phospholipid synthesis, 381*d*
protease, 238, 429
in sphinganine, 160
structure of, 54*t*
Serotonin, 455
Serum albumin, molecular weight, 78, 83*t*
Sex:
cell, 14
chromosome, 15
Shine-Dalgarno (S-D) sequence, 503
Shuttle:
glycerol 3-phosphate, 333*d*
malate-aspartate, 333*d*
Sialic acid, 42*s*, 161
Sickle-cell:
anemia, 18, 148
hemoglobin, 147, 439*d*
hydrophobic interactions in, 148
Sigmoidal binding curves:
2,3-bisphosphoglycerate (BPG), 119
binding to hemoglobin, 119, 120*d*
hemoglobin, 119*d*
structure, 119*s*
Sigmoidal curves, 271
Sigmoidal enzyme kinetics, 265, 267*d*
Sigmoidal plot, 266*d*, 267*d*, 273*d*
Signal peptidase, 506
Silkworm, cytochrome *c*, 98*dt*
Simian virus 40, SV40, genome size, 216*t*
Simple diffusion, across membranes, 177, 178*d*
Singer-Nicolson, model of membrane, 173*d*
Single-stranded DNA binding protein (SSB), 468, 469
Site-directed mutagenesis, 238
of enzyme, 239
Skeletal muscle, 17
Skin, 17, 121*t*, 122*t*, 124, 145, 149
Sliding contacts, hemoglobin, 148
Sliding filament model, of myofilament, 137
Small nuclear ribonucleoprotein, (snRNP), 497
Smith, G. D., 244, 245
Smooth endoplasmic reticulum (SER), 8*d*, 19
staining, 19
Smooth muscle, 17
snRNP, small nuclear ribonucleoprotein, 497
Sodium dodecyl sulfate (SDS), 83, 105
Solution:
crowding in, 114
ideal, 114
nonideal, 114
Somatic cell, 14, 15
Somatostatin, 427
Sonication, 19
cell rupturing, 19
vesicle preparation using, 170
Sørensen, S. P. L., 57
Species-selective toxicity, 245
Spectrin, 136, 137*d*, 150
α-helices, 150
coiled-coil, 150
cytoskeletal, 150
as part of cytoskeleton, 149
in cytoskeleton, 137*d*
heptad repeat, 150
hexamer, 150*d*
hydrophobic interactions in, 150
self-association, 150
tetramer, 135
Sperm, flagella, 141
Spermatozoon, 1
Sphinganine, 160*s*
Sphingolipid, 381
structure, 154, 160
Sphingolipidosis, 21
Sphingomyelin, 161*s*, 381, 382*d*, 383*d*
Sphingomyelinase, 381, 383*d*
Sphingosine, 160*s*, 382*d*, 383*d*
Spindle pole of MOTC, 143*d*
Spleen, 21
Spliceosome, 497
Splicing of RNA transcripts, 509
Squalene, 388*s*, 389*d*, 390*d*, 391*d*
epoxide, 391*d*
monooxygenase (2,3-epoxidizing), 391*d*
structure, 153*s*
synthesis from mevalonic acid, 182*s*
SSB, single-stranded binding protein, 469
Standard free energy, 293
Standard redox potential, 68
Standard state, 293
Starch, 46, 47
grains, 228
Starvation, ketone bodies in, 394*d*
State function, 290
Statistical mechanics, 290
Steady-state, 256, 259
in actin polymerization, 133
kinetic parameters, 261

Steady-state (*Cont.*):
rate equations, 257
Stearic acid, 155*s*
Stem cell, 1
Stercobilin, 453
Stercobilinogen, 453
Stereoisomer, 241
Steroid:
acetate as precursor, 154
coordinate regulation of synthesis, 390
definition, 163
effect on DNA transcription, 509
hormone, 390, 392, 393*d*, 399
Sterol, 362
definition, 163, 164
regulatory element-1, 390
Stop codon, 491, 504
Strain, molecular, 233
Strained conformation, 234
Streptolydigin, 500
Structural hierarchy in cells, 17, 22
Stylized representations of protein structures, 94
Subcellular organelle, 6*d*, 19
Submembrane cytoskeleton, 15
Substrate, 11*t*, 483
binding, 238
channeling, 442
concentration, 252
cycle, 301*d*, 308
Substrate binding affinity, 235, 236
Substrate-level phosphorylation:
in the citric acid cycle, 347*d*, 348
definition, 311, 402
in glycolysis, 315, 326, 402
Subtilisin, 238
Succinate:
in the citric acid cycle, 348*s*
in the glyoxylase cycle, 356*s*
product of prolylhydroxylase, 123*d*
Succinate dehydrogenase, 13*t*
in the citric acid cycle, 344*d*
location, 13*t*
oxaloacetate inhibition by, 351
Succinate/CoQ oxidoreductase, 405*t*
Succinyl-CoA, 451
in the citric acid cycle, 347*s*
synthetase, 347*s*
Sucrase, 46*s*, 330, 343
Sugar, 25
Suicide:
inhibitor, 445
substrate, 250
Sulfanilamide, 457*s*
Sulfate:
addition in chondroitin sulfate, 149
esters of sugars, 42
Sulfotransferase, 149
Sumner, J. B., 228
Supercoiled DNA, 217*d*
negative-, 217
positive-, 217
relaxation of, 217
supertwist, 217
Supramolecular structures, 108
assembly, 109
self-assembly, 104
Supravital dye, 12, 23
Svedberg, T., 81
Sweat gland, 17
Symbiont, intracellular, 12
Symbiotic bacteria, 420
Syn_psis, specialized form of exocytosis, 131
Synthesis of:
fatty acids, 374
glycogen, 327
palmitic acid, 374
phosphatidylcholine, 380*d*
triacylglycerol, 378*d*
1,2,3-triacylglycerol, 380*d*
unsaturated fatty acids, 376, 377*d*

T-lymphocyte, 456
Tactoid, 18
Tadpole, 455
Talose, 28*s*
Taq, DNA polymerase, 470, 478, 479
TATA box, Hogness box, 494, 496
Taurine, in bile salts, 168*s*
Taurocholate, 168*s*, 395
Tautomer, purine, 200
Tautomeric forms of bases, 469
Tay-Sachs, disease, 383*d*, 397
Tay-Sachs, ganglioside, 162*s*
Tears, 245
Teflon plunger, homogenizer, 5
Telomerase, 472*d*, 473, 486
Telomere, 472, 473, 498
Temperature-sensitive mutations, 470
Template, 459, 483
antisense strand, 489
DNA, 458, 465*d*, 481
in virus coat assembly, 110
Tendon, 121*t*, 122*t*, 124
Tense structure, hemoglobin, 121, 148
Ter, reaction, 259
Ter Quad, enzymatic reaction, 259*d*
TerA, *B*, *C*, *E* and *F*, 471
TerA-F, 471
Termination:
of chromosome replication, 460, 470, 471*d*
codon, 490*t*
of replication in eukaryotes, 486
Terminator protein, 471
Terminator sequences, 471
Terminus region of chromosome, 470, 471*d*
Terminus region of DNA, 460
Termite, 344, 420
Terpene:
acetate as precursor, 154
β-carotene, 163*s*
definition, 162
dolichol, 163*s*
farnesol, 162*s*
geraniol, 162*s*
limonene, 162*s*
ubiquinone, 163*s*
vitamin A, 162*s*
Tertiary lysosome, 22*d*
in macrophage, 22*d*
Testosterone, structure, 164*s*
Tetrahedral angle, 35
Tetrahedral intermediate, enzyme, 238
Tetrahedral intermediates, in enzymatic reactions, 232
Tetrahydrobiopterin, 425*d*, 426*s*
Tetrahydrofolate (THF), 447
Tetrahydrofolic acid, 448*st*
Tetrahymena thermophila, 497
Tetrameric protein, 23, 270
hemoglobin, 23
Tetranucleotide, 222*s*
Tetrapyrrole, in heme, 404*s*
Tetrasaccharide, 247
Tetrose, 25, 28
2-Thenoyltrifluoroacetone, inhibition of electron transport chain, 406*d*
Theorell-Chance, enzyme mechanism, 289
Thermal cycler, 479
Thermodynamic stability, 237
Thermodynamics, definition, 290
Thermostable DNA polymerase, 480
Thermus aquaticus, 478, 480
Thiamine, vitamin B_1, 352, 353*d*
Thiamine pyrophosphate (TPP), in pyruvate dehydrogenase, 116, 352, 353*d*
Thick filament, 139
myosin, 137
Thin filament, 131, 137, 138*d*, 139*d*
Thioether bonds, in cytochromes, 404
Thiogalactoside transacetylase, in *E. coli*, 508
Thiolase, 372*t*, 374, 388
Thioredoxin, 442*d*, 443
Thiosulfate, 229
Thiosulfate : cyanide sulfurtransferase, 229
Three-point attachment proposal, citrate binding to aconitase, 359*d*
Threonine, 423*t*
deaminase, 283
structure of, 54*t*
L-Threo-pentulose, 30
D-Threose, 27, 28
L-Threose, 28
Thromboxane, 386*ds*
[^{3}H]-Thymidine, 4
Thymidylate synthase, 443, 448
Thymine, 198, 199*s*, 205*s*, 458
dimer, 475*s*, 476*d*
Thymosin, sequestering of actin monomer, 134
Tissue, 3
Titin, 139
Tobacco mosaic virus (TMV), assembly, 109*d*, 110*d*
Toluidine blue, 18, 23
in histology, 18
Tomato bushy stunt virus, 114
Topoisomerase, 468, 471, 485
I, 468
II, 468
Topoisomers, 217, 468
Topology of DNA, 459
replication, 482
Total blood volume, 21
Toxicity, species-selective, 245
Trans elimination of H's, 348
Transacetylase, in pyruvate dehydrogenase, 116
Transaminase, 232
in malate-aspartate shuttle, 333

Transamination, 421
Transcription of DNA:
-10 region, 489, 492
bacterial, 511
complex, 495*d*
factors, 495*t*
Pribnow box, 492*d*
start site, 492
TFIIA, TFIIB, TFIID, TFIIE, 495*d*
Transcription factor, 494
tissue specific, 496
Transcriptional activation of OriC, 470
Transfer factor, heat stable, Ts, 503
Transfer RNA, 218
genes, 491
tRNA, 491, 500
Transferase, 230*t*
Transformation, of *E. coli*, 477
Transformylase, 503
Transition state of enzyme, 231, 232, 238, 243
Transition-state complex, 238, 249
Transition temperature:
cholesterol effect on, 174
of membrane lipids, 173
Transition-state analog, 236, 243, 244
Translation:
bacterial, 511
elongation, 503
initiation, 503
termination, 503
Translation of mRNA, 489
Translocase, 241
Transport:
amino acid, 430
dicarboxylate system, 431
L system, 431
L_y system, 431
Na^+/K^+ ATPase in, 430
Transport across membranes:
carrier mediated, 174
double reciprocal plot, 176*d*
facilitated, 177
kinetics of, 175*d*
pore, 174
saturation of, 176*d*
simple diffusion, 174
Transport protein, 174
Treadmilling, in tubulin, 140
α, α-Trehalose, 52
α, β-Trehalose, 50*s*
Triacylglycerol, 271, 362, 363*s*, 365, 366*s*, 374
in adipose tissue, 368
in depot fat, 158
digestion of, 363*d*
Fischer projection, 157*s*
sn- prefix, 157*s*
structure, 157*s*
synthesis of, 378*d*, 379*d*
1,2,3-Triacylglycerol, 378*d*
synthesis of, 380*d*
Tricarboxylic acid cycle, Krebs cycle, citric acid, 345, 346
Trifunctional enzyme, 438, 440
Triglyceride, 16, 362
storage vesicle, 16*d*
Trimolecular reaction, 274
Triose, 25
Triosephosphate isomerase, 314*d*
Triple helix, 123*d*
collagen, 122, 123*d*, 148
Triple-stranded helix, collagen, 121
Trisaccharide, 25, 45
Trisomy 21, 24
Tritiated amino acid, 4
tRNA, 240*d*
acceptor stem, 500, 501*d*
anticodon, 500, 501*d*
loops, 500, 501*d*
methionyl-, 503
transfer RNA, 491
types, 513, 514
Tropocollagen, 123, 124*d*
pentamicrofibril, 123
Tropoelastin, 125
Tropomodulin, 138, 139
Tropomyosin, 135, 136*d*, 137
coiled-coil, 135
cooperative binding of F-actin, 135
muscle, 135
Troponin C, 135, 136*d*, 137
pointed ends, 137
Troponin complex, 135
Trypsin, 102, 238, 248, 426, 428*d*
in protein sequencing, 80
Trypsin inhibitor, 426
Trypsinogen, 426, 427, 428*d*
Tryptophan, 423*t*
catabolism, 13*t*
location, 13*t*
structure of, 55*t*
Ts, heat stable transfer factor, 503
Tu, elongation factor, 503
Tubulin, 12, 113*d*, 131, 139, 150
critical concentration of monomers, 140
distribution in the cell, 140
dynamic instability, 140
GTP cap, 140
minus end, 140
organizing center, 140
plus end, 140
similarities to actin, 140
treadmilling, 140
Turnover number, enzyme, 256
Tus (terminus utilization substance), 471
Tus-Ter complex, 471
Tyrosine, 233, 244, 423*t*, 425*s*
structure of, 55*t*
L-Tyrosine, 240*d*
Tyrosyl-tRNA synthetase, 241

ΔU, internal energy change, 290, 292*d*
Ubiquinol, 404
Ubiquinone, 163*s*
coenzyme Q, 403, 404*s*
UDP-glucose-4-epimerase, 330
UDP-glucose pyrophosphorylase, 328
Ultracentrifugation, 4
Ultracentrifuge, 81
UMP-glucose, 328*s*
UMP synthase, 438
Uncompetitive inhibition, 254, 255*t*, 264, 280*d*
Uni, reaction, 259
Uni Bi enzymatic reaction, 259*d*
Uni Uni enzymatic reaction, 259*d*
Unidirectional replication, 461
Unitary rate constant, definition, 251
Unsaturated fatty acid, 155*s*
synthesis of, 376, 377*d*
Uracil, 198, 199*s*
Urate oxidase, 12
Urea, 17, 87*s*, 229*s*, 230*s*, 240*s*
cycle, 435, 437*d*
properties, 433
protein denaturation, 81
site of synthesis, 17
sources of nitrogen for, 434*d*
Urease (urea amidohydrolase), 228, 229, 241
molecular weight, 82*t*
nickel in, 241
sedimentation coefficient, 82*t*
Ureotele, 230
Uric acid, 456
Uricotele, 447
Uridine, 202
Urobilin, 453
Urobilinogen, 453
Uronate, 41*s*
Uronic acid, 41*s*, 126
Uroporphyrinogen III, 451*s*
Uroporphyrinogen cosynthase, 451
Uroporphyrinogen synthase, 451*d*
UV irradiation:
of DNA, 475
mutations and, 458
UvrABC, 476*d*

V_{max}, 253, 257, 261, 262*d*, 264*d*, 265*d*, 268, 276–278, 281, 282, 284, 285, 287–289
expression for, 260
Vacuole, 6*d*
Valine, 423*t*
structure of, 54*t*
Valinomycin:
diffusion potential, 417
effect on mitochondria, 409
mobile carrier, 181
potassium binding, 181*s*
structure, 181*s*
van der Waals bond, 247
Van't Hoff isochore, 295
Vasodilator, 398
Vector for DNA, 487
Vernier principle:
action of nebulin in actomyosin, 138
structural length control, 111*d*
Very low density lipoprotein (VLDL), 364
Vesicle, 9*d*, 141
fast axonal transport, 140, 141*d*
formation of, 170
structure, 169, 170*d*
Villin, 135, 136*d*
Vimentin, 145*t*
Virus:
assembly, 109*d*, 110*d*, 111*d*, 502*d*
coat protein, 110
DNA in, 199
proteases, 110
protein coat, 113
tomato bushy stunt, 114

Viscosity, high, 47
Vitamin:
  A, 162*s* 185*s*
  B_1, 352, 353*d*
  B_6, 421
  B_{12}, cobalamin, 449
  C, 406, 505
  D_3, 387, 399
  E, 153*s*
Vitamin C, and prolylhydroxylase, 123*d*
Vitreous body, 131
Vitreous humor, 122*t*
VLDL, 379
  definition, 168, 169*t*
  secretion of, 379*d*
Volume, total blood, 21

Water:
  ionic product of, 57
Water: (*Cont.*):
  ionization of, 57*d*
Watson, J. D., 207, 458
Wavelength, 3
Wax, 154
Wobble hypothesis, 501
Work, in thermodynamic experiments, 291
Wyman, J., 270

x-ray crystallography,
  of enzymes, 239
  of lysozyme, 245
  in protein structure determination, 99
Xanthine, 447*d*
Xanthine oxidase, 447*d*
Xeroderma pigmentosum, 476
Xylose, 28*s*, 45, 51
L-Xylulose, 31*s*
D-Xylulose 5-phosphate, 240*d*
L-Xylulose, 5-phosphate, 341*s*
Xylyltransferase, 149

Y chromosome, 15
Yeast, 419
  genome size, 216*t*

Z-disk, 137
  of actomyosin, 139
Zinc, in electrochemical cell, 296
Zn^{2+}:
  in carbonic anhydrase, 229
  in carboxypeptidase A, 248
  as electrophile, 249
Zymogen, 426, 428*d*