Dekker Encyclopedia of Nanoscience and Nanotechnology

Volume 4
Near–R
Pages 2677–3296

edited by

James A. Schwarz
Syracuse University, Syracuse, New York, U.S.A.

Cristian I. Contescu
Material Methods LLC, Newport Beach, California, U.S.A.

Karol Putyera
Shiva Technologies, Syracuse, New York, U.S.A.

MARCEL DEKKER, INC. NEW YORK • BASEL

Volume 1:
Main image: Courtesy of Tommaso Baldacchini and John T. Fourkas.
Detail: Courtesy of Kay Severin.

Volume 2:
Main image: Courtesy of Hamidou Haidara.
Detail: Courtesy of Jean-Yves Raty.

Volume 3:
Main image: Courtesy of Tony Van Buuren and Jürgen M. Plitzko.
Detail: Courtesy of Kay Severin.

Volume 4:
Main image: Courtesy of A. I. Gusev and colleagues.
Detail: Courtesy of Jean-Yves Raty.

Volume 5:
Main image: Courtesy of Ioan Balint.
Detail: Courtesy of Lars-Oliver Essen.

ISBN: Print: 0-8247-5055-1
ISBN: Online: 0-8247-5046-2
ISBN: Combo: 0-8247-4797-6

ISBN: Volume 1: 0-8247-5047-0
ISBN: Volume 2: 0-8247-5048-9
ISBN: Volume 3: 0-8247-5049-7
ISBN: Volume 4: 0-8247-5050-0
ISBN: Volume 5: 0-8247-5051-9

Library of Congress Cataloging-in-Publication Data
A catalog record of this book is available from the Library of Congress.

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-260-6300; fax: 41-61-260-6333

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/ Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

James A. Schwarz
Syracuse University, Syracuse, New York, U.S.A.
Cristian I. Contescu
Material Methods LLC, Newport Beach, California, U.S.A.
Karol Putyera
Shiva Technologies, Syracuse, New York, U.S.A.

Editorial Advisory Board

Laszlo B. Kish
Department of Electrical Engineering, Texas A&M University, College Station, Texas, U.S.A.

Kenneth Klabunde
Department of Chemistry, Kansas State University, Manhattan, Kansas, U.S.A.

James Leckie
Department of Civil and Environmental Engineering, School of Engineering, Stanford University, Stanford, California, U.S.A.

Charles M. Lieber
Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts, U.S.A.

Chad Mirkin
Department of Chemistry, Northwestern University, Evanston, Illinois, U.S.A.

Shuming Nie
Department of Chemistry, Indiana University, Bloomington, Indiana, U.S.A.

Jens K. Norskov
Department of Physics, Technical University of Denmark, Lyngby, Denmark

Mark Reed
Department of Electrical Engineering and Applied Physics, Yale University, New Haven, Connecticut, U.S.A.

David C. Rees
Astex Technology Ltd., Cambridge, U.K.

Stacey L. Ristinmaa-Sörensen
Department of Physics, University of Lund, Lund, Sweden

John A. Rogers
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey, U.S.A.

Debra R. Rolinson
Advanced Electrochemical Materials, Naval Research Laboratory, Washington, District of Columbia, U.S.A.

Jurgen Ruhe
Institute for Microsystem Technology, Albert-Ludwig University, Frieberg, Germany

Purnesh Seegopaul,
Umicore USA Inc., Hillsborough, New Jersey, U.S.A.

Stu Soled
ExxonMobil Research and Engineering, Annandale, New Jersey, U.S.A.

Stephan J. Stranick
National Institute of Standards and Technology (NIST), Gaithersburg, Maryland, U.S.A.

Arthur ten Wolde
Confederation of Netherlands Industry and Employers VNO-NCW, The Hague, The Netherlands

Frederick Tepper
Argonide Corp., Sanford, Florida, U.S.A.

Gary Tompa
Structured Materials Industries, Inc., Piscataway, New Jersey, U.S.A.

Robert Trew
Bradley Department of Electrical and Computer Engineering, Virginia Technical Institute, Blacksburg, Virginia, U.S.A.

Etienne F. Vansant
Department of Chemistry, University of Antwerp (UIA), Wilrijk, Belgium

Younan Xia
Department of Chemistry, University of Washington, Seattle, Washington, U.S.A.

Peidong Yang
Department of Chemistry, University of California, Berkeley, California, U.S.A.

List of Contributors

Nehal I. Abu-Lail / *Worcester Polytechnic Institute, Worcester, Massachusetts, U.S.A.*

R. Shane Addleman / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*

Rigoberto C. Advincula / *University of Alabama at Birmingham, Birmingham, Alabama, U.S.A.* / *University of Houston, Houston, Texas, U.S.A.*

Sean R. Agnew / *University of Virginia, Charlottesville, Virginia, U.S.A.*

Brahim Akdim / *AFRL/ML, Wright-Patterson Air Force Base, Ohio, U.S.A.*

Joseph Akkara / *National Science Foundation, Arlington, Virginia, U.S.A.*

Tomoyuki Akutagawa / *Hokkaido University, Sapporo, Japan*

Markus Albrecht / *Institut für Organische Chemie, Aachen, Germany*

A. Paul Alivisatos / *University of California, Berkeley, California, U.S.A.*

Charles W. Allen / *Argonne National Laboratory, Argonne, Illinois, U.S.A.*

Tsuneya Ando / *Tokyo Institute of Technology, Tokyo, Japan*

Ronald P. Andres / *Purdue University, West Lafayette, Indiana, U.S.A.*

Mikhail A. Anisimov / *University of Maryland, College Park, Maryland, U.S.A.*

Masakazu Aono / *National Institute for Materials Science, Ibaraki, Japan*

Hajime Asahi / *Osaka University, Osaka, Japan*

K. Asakawa / *The Femtosecond Technology Research Association (FESTA), Ibaraki, Japan*

B. J. Ash / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*

Masafumi Ata / *SONY Corporation, Yokohama, Japan*

Plamen B. Atanassov / *The University of New Mexico, Albuquerque, New Mexico, U.S.A.*

Elizabeth K. Auty / *University of Bristol, Bristol, United Kingdom*

Eric Ayars / *California State University—Chico, Chico, California, U.S.A.*

George Bachand / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*

Kannan Balasubramanian / *Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany*

Ganesh Balasubramanian / *University of Cincinnati, Cincinnati, Ohio, U.S.A.*

Anna C. Balazs / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*

Tommaso Baldacchini / *Boston College, Chestnut Hill, Massachusetts, U.S.A.*

Ioan Balint / *Romanian Academy, Bucharest, Romania*

S. Bandyopadhyay / *Virginia Commonwealth University, Richmond, Virginia, U.S.A.*

Rajarshi Banerjee / *The Ohio State University, Columbus, Ohio, U.S.A.*

Sarbajit Banerjee / *State University of New York at Stony Brook, Stony Brook, New York, U.S.A.*

Maria C. Bartelt / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*

James D. Batteas / *National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*

Jan Becher / *University of Southern Denmark, Odense, Denmark*

Paul D. Beer / *University of Oxford, Oxford, United Kingdom*

Silke Behrens / *Institute of Technical Chemistry, Karlsruhe, Germany*

Elena Bekyarova / *University of California, Riverside, California, U.S.A.*

G. Benedek / *Università di Milano Bicocca, Milan, Italy*
Yaakov Benenson / *Weizmann Institute of Science, Rehovot, Israel*
M. Bernasconi / *Università di Milano Bicocca, Milan, Italy*
Steven L. Bernasek / *Princeton University, Princeton, New Jersey, U.S.A.*
Alexey Bezryadin / *University of Illinois at Urbana-Champaign, Champaign, Illinois, U.S.A.*
Ajay Kumar Bhagi / *University of Delhi, Delhi, India*
Dhruba Jyoti Bharali / *State University of New York, Buffalo, New York, U.S.A.*
T. A. Birks / *University of Bath, Bath, United Kingdom*
Jerome C. Birnbaum / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Harvey W. Blanch / *University of California, Berkeley, California, U.S.A.*
Alexandre Blumstein / *University of Massachusetts, Lowell, Massachusetts, U.S.A.*
Andrew B. Bocarsly / *Princeton University, Princeton, New Jersey, U.S.A.*
Michael R. Bockstaller / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
A. Bogicevic / *Ford Motor Company, Dearborn, Michigan, U.S.A.*
Anthony Bollinger / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
Olga V. Boltalina / *M. V. Lomonosov Moscow State University, Moscow, Russia*
Mila Boncheva / *Harvard University, Cambridge, Massachusetts, U.S.A.*
Helmut Bönnemann / *Max Planck Institut für Kohlenforschung, Mülheim an der Ruhr, Germany*
Arijit Bose / *University of Rhode Island, Kingston, Rhode Island, U.S.A.*
Eduardo J. Bottani / *Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas, (UNLP. CIC, CONICET), La Plata, Argentina*
Laurent Bouteiller / *Université Pierre et Marie Curie, Paris, France*
Charles Brands / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*
Paul S. Braterman / *University of North Texas, Denton, Texas, U.S.A.*
Roberto J. Brea / *Universidade de Santiago de Compostela, Santiago de Compostela, Spain*
Lyudmila M. Bronstein / *Indiana University, Bloomington, Indiana, U.S.A.*
E. A. Martijn Brouwer / *University of Twente, Enschede, The Netherlands*
Ronald C. Brown / *Mercyhurst College, Erie, Pennsylvania, U.S.A.*
Nigel D. Browning / *University of California, Davis, California, U.S.A.*
Marko Burghard / *Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany*
Daryle H. Busch / *University of Kansas, Lawrence, Kansas, U.S.A.*
Harshala Butala / *University of Mississippi, University, Mississippi, U.S.A.*
Gavin A. Buxton / *University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.*
Mark E. Byrne / *Auburn University, Auburn, Alabama, U.S.A.*
Valérie Cabuil / *Université Pierre et Marie Curie, Paris, France*
Tahir Cader / *Isothermal Systems Research, Clarkston, Washington, U.S.A.*
Yuguang Cai / *Princeton University, Princeton, New Jersey, U.S.A.*
Terri A. Camesano / *Worcester Polytechnic Institute, Worcester, Massachusetts, U.S.A.*
Autumn T. Carlsen / *University at Albany—State University of New York, Albany, New York, U.S.A.*
Andrew D. W. Carswell / *University of Oklahoma, Norman, Oklahoma, U.S.A.*
Alessandro Casnati / *Università di Parma, Parma, Italy*
Mehmet S. Celik / *Istanbul Technical University, Istanbul, Turkey*
Gianfranco Cerofolini / *STMicroelectronics, Catania, Italy*
Neil R. Champness / *The University of Nottingham, Nottingham, United Kingdom*
Selena Chan / *University of Rochester, Rochester, New York, U.S.A.*
Christophe Chassenieux / *Université Pierre et Marie Curie, Paris, France*

James R. Chelikowsky / *University of Minnesota, Minneapolis, Minnesota, U.S.A.*

Bin Chen / *National Aeronautics and Space Administration (NASA), Moffett Field, California, U.S.A.*

Chenggang Chen / *University of Dayton Research Institute, Dayton, Ohio, U.S.A.* / *Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio, U.S.A*

Zhan Chen / *University of Michigan, Ann Arbor, Michigan, U.S.A.*

Lifeng Chi / *Westfälische Wilhelms-Universität Münster, Münster, Germany*

Russell R. Chianelli / *The University of Texas, El Paso, Texas, U.S.A.*

Jeong-Min Cho / *University of Florida, Gainesville, Florida, U.S.A.*

Kwang-Min Choi / *Osaka University, Osaka, Japan*

Thomas Clifford / *University of Kansas, Lawrence, Kansas, U.S.A.*

Luise S. Couchman / *Naval Research Laboratory, Washington, District of Columbia, U.S.A.*

Christophe Coudret / *Centre d'Elaboration de Matériaux et d'Etudes Structurales (CEMES), Centre National de la Recherche Scientifique (CNRS), Toulouse, France*

Alexander Couzis / *City College of City University of New York, New York, New York, U.S.A.*

Mercedes Crego-Calama / *University of Twente, Enschede, The Netherlands*

B. Roldan Cuenya / *University of California, Santa Barbara, California, U.S.A.*

Brian M. Cullum / *University of Maryland, Baltimore County, Baltimore, Maryland, U.S.A.*

Peter T. Cummings / *Vanderbilt University, Nashville, Tennessee, U.S.A.* / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*

Joan E. Curry / *University of Arizona, Tucson, Arizona, U.S.A.*

Lars Dähne / *Capsulution Nanoscience AG, Berlin, Germany*

Sheng Dai / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*

Enrico Dalcanale / *Università di Parma, Parma, Italy*

Anthony P. Davis / *University of Bristol, Bristol, United Kingdom*

Jason J. Davis / *University of Oxford, Oxford, United Kingdom*

Richey M. Davis / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*

Robert J. Davis / *University of Virginia, Charlottesville, Virginia, U.S.A.*

Thomas F. Degnan, Jr. / *ExxonMobil Research and Engineering Company, Annandale, New Jersey, U.S.A.*

Thierry Delair / *École Normale Supérieure de Lyon, Lyon, France*

Aránzazu del Campo / *Max Planck Institute for Polymer Research, Mainz, Germany*

Pedro A. Derosa / *University of South Carolina, Columbia, South Carolina, U.S.A.*

Roland Dersch / *Philipps-Universität Marburg, Marburg, Germany*

A. Prasanna de Silva / *School of Chemistry at Queen's University, Belfast, Northern Ireland*

James J. De Yoreo / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*

Stephan Diekmann / *Institut für Molekulare Biotechnologie, Jena, Germany*

João C. Diniz da Costa / *The University of Queensland, Brisbane, Queensland, Australia*

Dionysios D. Dionysiou / *University of Cincinnati, Cincinnati, Ohio, U.S.A.*

Mark M. Disko / *ExxonMobil Research and Engineering, Annandale, New Jersey, U.S.A.*

Stephen K. Doorn / *Los Alamos National Laboratory, Los Alamos, New Mexico, U.S.A.*

Bogdan Dragnea / *Indiana University, Bloomington, Indiana, U.S.A.*

Charles Michael Drain / *Hunter College of the City University of New York, New York, New York, U.S.A.* / *The Rockefeller University, New York, New York, U.S.A.*

Jaroslaw Drelich / *Michigan Technological University, Houghton, Michigan, U.S.A.*

G. Dresselhaus / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*

Mildred S. Dresselhaus / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*

Xiaofeng Duan / *AFRL/ML, Wright-Patterson Air Force Base, Ohio, U.S.A.*

Yves F. Dufrêne / *Université catholique de Louvain, Louvain-la-Neuve, Belgium*

Etienne Duguet / *Institut de Chimie de la Matière Condensée de Bordeaux, CNRS, Pessac, France*
Kathleen A. Dunn / *University at Albany—State University of New York, Albany, New York, U.S.A.*
R. D. Dupuis / *The University of Texas at Austin, Austin, Texas, U.S.A.*
Anne-Sophie Duwez / *Université catholique de Louvain, Louvain-la-Neuve, Belgium*
Kohki Ebitani / *Osaka University, Osaka, Japan*
A. Eitan / *Rensselaer Polytechnic Institute, Troy, New York, U.S.A.*
Jeremiah Ejiofor / *Purdue University, West Lafayette, Indiana, U.S.A.*
Vladimir I. Elokhin / *Boreskov Institute of Catalysis, Novosibirsk, Russia*
Todd Emrick / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*
Morinobu Endo / *Shinshu University, Nagano, Japan*
Jonah Erlebacher / *Johns Hopkins University, Baltimore, Maryland, U.S.A.*
Bahri Ersoy / *Afyon Kocatepe University, Afyon, Turkey*
Lars-Oliver Essen / *Philipps University, Marburg, Germany*
Louise S. Evans / *University of Southampton, Southampton, United Kingdom*
Jeffrey D. Evanseck / *Duquesne University, Pittsburgh, Pennsylvania, U.S.A.*
Rodney C. Ewing / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
D. Howard Fairbrother / *The Johns Hopkins University, Baltimore, Maryland, U.S.A.*
Xiaowu Fan / *University of Alabama at Birmingham, Birmingham, Alabama, U.S.A.*
Zhigang Fang / *University of Utah, Salt Lake City, Utah, U.S.A.*
Philippe M. Fauchet / *University of Rochester, Rochester, New York, U.S.A.*
H.-J. Fecht / *University of Ulm, Ulm, Germany*
X. Feng / *Ferro Corporation, Independence, Ohio, U.S.A.*
Ben L. Feringa / *University of Groningen, Groningen, The Netherlands*
Shaun F. Filocamo / *Boston University, Boston, Massachusetts, U.S.A.*
Paul D. I. Fletcher / *University of Hull, Hull, United Kingdom*
Robert E. Fontana, Jr. / *IBM Almaden Research Center, San Jose, California, U.S.A.*
Christy Ford / *Tulane University, New Orleans, Louisiana, U.S.A.*
John T. Fourkas / *Boston College, Chestnut Hill, Massachusetts, U.S.A.*
Elzbieta Frackowiak / *Poznan University of Technology, Poznan, Poland*
Thomas Franklin / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
Hamish L. Fraser / *The Ohio State University, Columbus, Ohio, U.S.A.*
Wolfgang Fritzsche / *Institute for Physical High Technology (IPHT), Jena, Germany*
Glen E. Fryxell / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Jürgen-Hinrich Fuhrhop / *Freie Universität Berlin, Berlin, Germany*
T. A. Fulton / *Lucent Technologies, Murray Hill, New Jersey, U.S.A.*
Kunio Furusawa / *University of Tsukuba, Ibaraki, Japan*
Charles-André Fustin / *Max Planck Institute for Polymer Research, Mainz, Germany*
Aurelian C. Gâlcă / *University of Twente, Enschede, The Netherlands*
Philip A. Gale / *University of Southampton, Southampton, United Kingdom*
Giulia Galli / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Ignacio L. Garzón / *Universidad Nacional Autónoma de México, Mexico City, Mexico*
J. Gaudioso / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*
Andrew J. Gellman / *Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.*
Bruce C. Gibb / *University of New Orleans, New Orleans, Louisiana, U.S.A.*
Jeffrey C. Gibeling / *University of California, Davis, California, U.S.A.*
Harry W. Gibson / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*

David S. Ginley / *National Renewable Energy Laboratory (NREL), Golden, Colorado, U.S.A.*

Anit Giri / *Nanomat, Inc., North Huntingdon, Pennsylvania, U.S.A.*

Michael Gleiche / *Westfälische Wilhelms-Universität Münster, Münster, Germany*

William A. Goddard / *California Institute of Technology, Pasadena, California, U.S.A.*

George W. Gokel / *Washington University School of Medicine, St. Louis, Missouri, U.S.A.*

Shun-ichi Gonda / *Fukui University of Technology, Fukui, Japan*

D. Wayne Goodman / *Texas A&M University, College Station, Texas, U.S.A.*

Lionel Goodman / *Rutgers, The State University of New Jersey, New Brunswick, New Jersey, U.S.A.*

Brian P. Grady / *University of Oklahoma, Norman, Oklahoma, U.S.A.*

Juan R. Granja / *Universidade de Santiago de Compostela, Santiago de Compostela, Spain*

Christine S. Grant / *North Carolina State University, Raleigh, North Carolina, U.S.A.*

Vicki H. Grassian / *University of Iowa, Iowa City, Iowa, U.S.A.*

Jay W. Grate / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*

Elias Greenbaum / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.* / *The University of Tennessee, Knoxville, Tennessee, U.S.A.*

Andreas Greiner / *Philipps-Universität Marburg, Marburg, Germany*

Mark W. Grinstaff / *Boston University, Boston, Massachusetts, U.S.A.*

Edward Grochowski / *IBM Almaden Research Center, San Jose, California, U.S.A.*

Gwen M. Gross / *University of Washington, Seattle, Washington, U.S.A.*

Joanna R. Groza / *University of California, Davis, California, U.S.A.*

Dirk M. Guldi / *University of Notre Dame, Notre Dame, Indiana, U.S.A.*

Jianchang Guo / *Emory University, Atlanta, Georgia, U.S.A.*

Vinay K. Gupta / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*

Alexandr I. Gusev / *Russian Academy of Sciences, Yekaterinburg, Russia*

Matthew Guzy / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*

Kwon-Soo Ha / *Kangwon National University School of Medicine, Chunchon, South Korea*

Robert C. Haddon / *University of California, Riverside, California, U.S.A.*

Hamidou Haidara / *Institut de Chimie des Surfaces et Interfaces-ICSI-CNRS, Mulhouse, France*

Wolfgang Haiss / *University of Liverpool, Liverpool, United Kingdom*

Kimberly Hamad-Schifferli / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*

Jingmin Han / *University of Alberta, Edmonton, Alberta, Canada*

Encai Hao / *Northwestern University, Evanston, Illinois, U.S.A.*

J. W. Harrell / *The University of Alabama, Tuscaloosa, Alabama, U.S.A.*

Stephen J. Haswell / *University of Hull, Hull, United Kingdom*

Bryan C. Hathorn / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*

Takuya Hayashi / *Shinshu University, Nagano, Japan*

Jibao He / *Tulane University, New Orleans, Louisiana, U.S.A.*

J. R. Heflin / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*

James M. Helt / *College of Staten Island and Graduate Center of the City University of New York, New York, New York, U.S.A.*

Hans-Peter Hentze / *University of Delaware, Newark, Delaware, U.S.A.*

Henry Hess / *University of Washington, Seattle, Washington, U.S.A.*

Daniel A. Higgins / *Kansas State University, Manhattan, Kansas, U.S.A.*

Alexander Hillisch / *EnTec GmbH, Jena, Germany*

J. Zachary Hilt / *The University of Texas, Austin, Texas, U.S.A.*

Satoshi Hirosawa / *Sumitomo Special Metals Company, Ltd., Osaka, Japan*

Andreas Hirsch / *Friedrich Alexander Universität Erlangen-Nürnberg, Erlangen, Germany*
Eric M. V. Hoek / *University of California, Riverside, California, U.S.A.*
Peter M. Hoffmann / *Wayne State University, Detroit, Michigan, U.S.A.*
J. Hone / *Columbia University, New York, New York, U.S.A.*
Louisa J. Hope-Weeks / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
David Hopkins / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
Scott R. Horner / *University of Rochester, Rochester, New York, U.S.A.*
Yanwen Hou / *Kansas State University, Manhattan, Kansas, U.S.A.*
Hui Hu / *University of California, Riverside, California, U.S.A.*
Zhibing Hu / *University of North Texas, Denton, Texas, U.S.A.*
Yu-Wen Huang / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
Mark Hughes / *University of Cambridge, Cambridge, United Kingdom*
Deborah E. Hunka / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*
Ho Jung Hwang / *Chung-Ang University, Seoul, South Korea*
Akihisa Inoue / *Tohoku University, Sendai, Japan*
Yasushi Inouye / *Osaka University, Osaka, Japan*
Stephan Irle / *Emory University, Atlanta, Georgia, U.S.A.*
Kelvin Isaacson / *University of Alberta, Edmonton, Alberta, Canada*
Toshihiro Ishikawa / *Ube Industries Ltd., Ube Research Laboratory, Yamaguchi, Japan*
Hiroyuki Isobe / *The University of Tokyo, Tokyo, Japan*
Mikhail E. Itkis / *University of California, Riverside, California, U.S.A.*
Takashi Ito / *Fujitsu Ltd., Tokyo, Japan*
Heinrich M. Jaeger / *University of Chicago, Chicago, Illinois, U.S.A.*
Dustin K. James / *Rice University, Houston, Texas, U.S.A.*
David B. Janes / *Purdue University, West Lafayette, Indiana, U.S.A.*
Andrés Jaramillo-Botero / *Pontificia Universidad Javeriana, Cali, Colombia*
Anna Jawor / *University of California, Riverside, California, U.S.A.*
Andreas Jentys / *Technische Universität München, Garching, Germany*
Christian Joachim / *Centre d'Elaboration de Matériaux et d'Etudes Structurales (CEMES), Centre National de la Recherche Scientifique (CNRS), Toulouse, France*
Vijay John / *Tulane University, New Orleans, Louisiana, U.S.A.*
Katrina A. Jolliffe / *University of Sydney, Sydney, New South Wales, Australia*
Ulrich Jonas / *Max Planck Institute for Polymer Research, Mainz, Germany*
Clinton D. Jones / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
A. Jorio / *Universidade Federal de Minas Gerais, Belo Horizonte, Brazil*
Leonid Kaledin / *Argonide Corporation, Sanford, Florida, U.S.A.*
Alain E. Kaloyeros / *University at Albany—State University of New York, Albany, New York, U.S.A.*
Toshiya Kamikado / *Communications Research Laboratory, Kobe, Japan*
Kiyotomi Kaneda / *Osaka University, Osaka, Japan*
Jeong Won Kang / *Chung-Ang University, Seoul, South Korea*
David Kaplan / *Tufts University, Medford, Massachusetts, U.S.A.*
Pramesh N. Kapoor / *University of Delhi, Delhi, India*
Hiromichi Kataura / *Tokyo Metropolitan University, Tokyo, Japan*
Satoshi Kawata / *Osaka University, Osaka, Japan*
Christine D. Keating / *Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
Klaus Kern / *Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany*

Miklos Kertesz / *Georgetown University, Washington, District of Columbia, U.S.A.*
Abbas Khaleel / *United Arab Emirates University, Al-Ain, United Arab Emirates*
Aleksandr Khitun / *University of California, Los Angeles, California, U.S.A.*
David T. Kim / *University of California, Berkeley, California, U.S.A.*
Dong Wook Kim / *Korea Research Institute of Chemical Technology, Daejeon, South Korea*
Jungbae Kim / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Kwang S. Kim / *Pohang University of Science and Technology, Pohang, South Korea*
Sungsoo Kim / *University of Illinois at Urbana-Champaign, Champaign, Illinois, U.S.A.*
Yoong Ahm Kim / *Shinshu University, Nagano, Japan*
R. Allen Kimel / *Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
Laszlo B. Kish / *Texas A&M University, College Station, Texas, U.S.A.*
Kenneth J. Klabunde / *Kansas State University, Manhattan, Kansas, U.S.A.*
Jennifer E. Klare / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Paul G. Klemens / *University of Connecticut, Storrs, Connecticut, U.S.A.*
Robert F. Klie / *Brookhaven National Laboratory, Upton, New York, U.S.A.*
J. C. Knight / *University of Bath, Bath, United Kingdom*
Carl C. Koch / *North Carolina State University, Raleigh, North Carolina, U.S.A.*
Kenichiro Koga / *Okayama University, Okayama, Japan*
Punit Kohli / *University of Florida, Gainesville, Florida, U.S.A.*
S. Kohmoto / *The Femtosecond Technology Research Association (FESTA), Ibaraki, Japan*
Toshihiro Kondo / *Hokkaido University, Sapporo, Japan*
E. Stefan Kooij / *University of Twente, Enschede, The Netherlands*
Nicholas A. Kotov / *Oklahoma State University, Stillwater, Oklahoma, U.S.A.*
Vladimír Král / *Institute of Chemical Technology, Prague, Czech Republic*
Ilona Kretzschmar / *Yale University, New Haven, Connecticut, U.S.A.*
Xiomara C. Kretschmer / *The University of Texas, El Paso, Texas, U.S.A.*
A. J. Kropf / *Argonne National Laboratory, Argonne, Illinois, U.S.A.*
Bon-Cheol Ku / *University of Massachusetts, Lowell, Massachusetts, U.S.A.*
Clifford P. Kubiak / *University of California—San Diego, La Jolla, California, U.S.A.*
Debasish Kuila / *Louisiana Tech University, Ruston, Louisiana, U.S.A.*
Jayant Kumar / *University of Massachusetts, Lowell, Massachusetts, U.S.A.*
Nitin Kumar / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
Daniel Y. Kwok / *University of Alberta, Edmonton, Alberta, Canada*
Soon-Bark Kwon / *Kwangju Institute of Science and Technology, Gwangju, South Korea*
Dominique Larcher / *Université de Picardie—Jules Verne, Amiens, France*
Sarah C. Larsen / *University of Iowa, Iowa City, Iowa, U.S.A.*
Jean-Pierre Launay / *Centre d'Elaboration de Matériaux et d'Etudes Structurales (CEMES), Centre National de la Recherche Scientifique (CNRS), Toulouse, France*
Valentina Lazarescu / *Institute of Physical Chemistry "I. G. Murgulescu," Bucharest, Romania*
Nikolai N. Ledentsov / *Technische Universität Berlin, Berlin, Germany*
Haiwon Lee / *Hanyang University, Seoul, South Korea*
Han Myoung Lee / *Pohang University of Science and Technology, Pohang, South Korea*
Ken Won Lee / *Kwangju Institute of Science and Technology, Gwangju, South Korea*
Seunghyun Lee / *Hanyang University, Seoul, South Korea*
Steven Lenhert / *Westfälische Wilhelms-Universität Münster, Münster, Germany*
Kam W. Leong / *The Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.*

Johannes A. Lercher / *Technische Universität München, Garching, Germany*
Marat I. Lerner / *Russian Academy of Sciences, Tomsk, Russia*
Rastislav Levicky / *Columbia University, New York, New York, U.S.A.*
Guangtao Li / *Freie Universität Berlin, Berlin, Germany*
Liang-shi Li / *University of California, Berkeley, California, U.S.A.*
Tianquan Lian / *Emory University, Atlanta, Georgia, U.S.A.*
Chengdu Liang / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*
Xuemei Liang / *Wayne State University, Detroit, Michigan, U.S.A.*
Yeshayahu Lifshitz / *City University of Hong Kong, Hong Kong, China / Soreq Nuclear Research Center, Yavne, Israel*
Peter T. Lillehei / *NASA-Langley Research Center, Hampton, Virginia, U.S.A.*
Xiao Min Lin / *Argonne National Laboratory, Argonne, Illinois, U.S.A.*
Yuehe Lin / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Leonard F. Lindoy / *University of Sydney, Sydney, New South Wales, Australia*
J. Liu / *Sandia National Laboratory, Albuquerque, New Mexico, U.S.A.*
Jian Liu / *Michigan Technological University, Houghton, Michigan, U.S.A.*
Jianlin Liu / *University of California, Riverside, California, U.S.A.*
Jingyue Liu / *Monsanto Company, St. Louis, Missouri, U.S.A.*
Zhen Liu / *Columbia University, New York, New York, U.S.A.*
Jason Locklin / *University of Houston, Houston, Texas, U.S.A.*
Yueh-Lin Loo / *The University of Texas, Austin, Texas, U.S.A.*
Gabriel P. López / *The University of New Mexico, Albuquerque, New Mexico, U.S.A.*
Dmitri Valentinovich Louzguine / *Tohoku University, Sendai, Japan*
J. Christopher Love / *Harvard University, Cambridge, Massachusetts, U.S.A.*
Fang Lu / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
G. Q. (Max) Lu / *The University of Queensland, Brisbane, Queensland, Australia*
Nan Lu / *Westfälische Wilhelms-Universität Münster, Münster, Germany*
Yan-Yeung Luk / *University of Wisconsin—Madison, Madison, Wisconsin, U.S.A.*
Simon O. Lumsdon / *DuPont Central Research and Development, Wilmington, Delaware, U.S.A.*
L. Andrew Lyon / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
Alexander P. Lyubartsev / *Stockholm University, Stockholm, Sweden*
Jeffry D. Madura / *Duquesne University, Pittsburgh, Pennsylvania, U.S.A.*
Kenzo Maehashi / *Osaka University, Osaka, Japan*
Thomas F. Magnera / *University of Colorado, Boulder, Colorado, U.S.A.*
Amitesh Maiti / *Accelrys Inc., San Diego, California, U.S.A.*
Amarnath Maitra / *University of Delhi, Delhi, India*
Tatiana L. Makarova / *Umea University, Umea, Sweden*
Sergey A. Maksimenko / *Belarus State University, Minsk, Belarus*
Arif A. Mamedov / *Oklahoma State University and Nomadics, Inc., Stillwater, Oklahoma, U.S.A.*
Nataliya N. Mamedova / *Oklahoma State University, Stillwater, Oklahoma, U.S.A.*
Guangzhao Mao / *Wayne State University, Detroit, Michigan, U.S.A.*
Gabriela Marinescu / *"I. G. Murgulescu" Institute of the Romanian Academy, Bucharest, Romania*
C. L. Marshall / *Argonne National Laboratory, Argonne, Illinois, U.S.A.*
Charles R. Martin / *University of Florida, Gainesville, Florida, U.S.A.*
Shinro Mashiko / *Communications Research Laboratory, Kobe, Japan*
Hiroshi Matsui / *The City University of New York, Hunter College, New York, New York, U.S.A.*

Hideo Matsumura / *National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan*
Hideki Matsuoka / *Kyoto University, Kyoto, Japan*
S. V. Mattigod / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
F. Maurigh / *DIEGM University of Udine, Udine, Italy*
Mikhail V. Maximov / *Ioffe Physico-Technical Institute, St. Petersburg, Russia*
Clare McCabe / *Colorado School of Mines, Golden, Colorado, U.S.A.*
Nathan D. McClenaghan / *School of Chemistry at Queen's University, Belfast, Northern Ireland*
E. W. McFarland / *University of California, Santa Barbara, California, U.S.A.*
Duncan J. McGillivray / *Oxford University, Oxford, United Kingdom*
Gary McPherson / *Tulane University, New Orleans, Louisiana, U.S.A.*
Claude F. Meares / *University of California, Davis, California, U.S.A.*
Mark S. Meier / *University of Kentucky, Lexington, Kentucky, U.S.A.*
Latika Menon / *Texas Tech University, Lubbock, Texas, U.S.A.*
Curtis W. Meuse / *National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.*
Agnes A. Mewe / *University of Twente, Enschede, The Netherlands*
F. Miani / *DIEGM University of Udine, Udine, Italy*
Weifang Miao / *Nanomat, Inc., North Huntingdon, Pennsylvania, U.S.A.*
Josef Michl / *University of Colorado, Boulder, Colorado, U.S.A.*
Laszlo Mihaly / *The State University of New York at Stony Brook, Stony Brook, New York, U.S.A.*
Abigail E. Miller / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Benjamin L. Miller / *University of Rochester, Rochester, New York, U.S.A.*
J. T. Miller / *British Petroleum (BP), Naperville, Illinois, U.S.A.*
Scott A. Miller / *University of Florida, Gainesville, Florida, U.S.A.*
Ioana Mindru / *"I. G. Murgulescu" Institute of the Romanian Academy, Bucharest, Romania*
Christophe Mingotaud / *Laboratoire des IMRCP, Université Paul Sabatier, Toulouse, France*
John W. Mintmire / *Oklahoma State University, Stillwater, Oklahoma, U.S.A.*
Francis M. Mirabella, Jr. / *Equistar Chemicals, LP, Cincinnati, Ohio, U.S.A.*
Susmita Mitra / *University of Delhi, Delhi, India*
Akane Miyazaki / *Tokyo Institute of Technology, Yokohama, Japan*
Tomoo Mizugaki / *Osaka University, Osaka, Japan*
Peter Moeck / *Portland State University, Portland, Oregon, U.S.A.*
Carlo D. Montemagno / *University of California, Los Angeles, California, U.S.A.*
Keiji Morokuma / *Emory University, Atlanta, Georgia, U.S.A.*
Karine Mougin / *Institut de Chimie des Surfaces et Interfaces-ICSI-CNRS, Mulhouse, France*
Emiko Mouri / *Kyoto University, Kyoto, Japan*
Douglas Mulhall / *Marsh Harbour, Bahamas*
Ravichandra S. Mulukutla / *Nanoscale Materials, Inc., Manhattan, Kansas, U.S.A.*
Kei Murakoshi / *Osaka University, Osaka, Japan*
Makoto Murata / *Osaka University, Osaka, Japan*
Michael Murphey / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
Mustafa Musameh / *New Mexico State University, Las Cruces, New Mexico, U.S.A.*
Aleksandr V. Myshlyavtsev / *Omsk State Technical University, Omsk, Russia*
Edward M. Nadgorny / *Michigan Technological University, Houghton, Michigan, U.S.A.*
K. S. Nagabhushana / *Max Planck Institut für Kohlenforschung, Mülheim an der Ruhr, Germany*
Eiichi Nakamura / *The University of Tokyo, Tokyo, Japan*
H. Nakamura / *The Femtosecond Technology Research Association (FESTA), Ibaraki, Japan*

Takayoshi Nakamura / *Hokkaido University, Sapporo, Japan*
Hisao Nakashima / *Osaka University, Osaka, Japan*
Yoshihiro Nakato / *Osaka University, Osaka, Japan*
Richard M. Narske / *Augustana College, Rock Island, Illinois, U.S.A.*
Alexandra Navrotsky / *University of California, Davis, California, U.S.A.*
David A. Nelson / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Sarah A. Nespolo / *University of Melbourne, Melbourne, Victoria, Australia*
M. K. Neylon / *Argonne National Laboratory, Argonne, Illinois, U.S.A.*
Patrick J. Neyman / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*
K. Y. Simon Ng / *Wayne State University, Detroit, Michigan, U.S.A.*
Richard J. Nichols / *University of Liverpool, Liverpool, United Kingdom*
Marc Nicklaus / *National Cancer Institute, Frederick, Maryland, U.S.A.*
David E. Nikles / *The University of Alabama, Tuscaloosa, Alabama, U.S.A.*
S. Nishikawa / *The Femtosecond Technology Research Association (FESTA), Ibaraki, Japan*
Sandip Niyogi / *University of California, Riverside, California, U.S.A.*
Ijeoma M. Nnebe / *Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.*
Donald W. Noid / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*
Tyler B. Norsten / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*
Aleksandr Noy / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Stacie Nunes / *State University of New York at New Paltz, New Paltz, New York, U.S.A.*
Masayuki Okamura / *Hokkaido University, Sapporo, Japan*
Yuji Okawa / *National Institute for Materials Science, Ibaraki, Japan*
Masahiko Ooe / *Osaka University, Osaka, Japan*
Neil S. Oxtoby / *The University of Nottingham, Nottingham, United Kingdom*
Ruth Pachter / *AFRL/ML, Wright-Patterson Air Force Base, Ohio, U.S.A.*
Christopher S. Palenik / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
Chandra Sekhar Palla / *University of Delaware, Newark, Delaware, U.S.A.*
Raghuveer Parthasarathy / *University of Chicago, Chicago, Illinois, U.S.A.* / *University of California, Berkeley, California, U.S.A.*
Sandeep Patel / *Hunter College of the City University of New York, New York, New York, U.S.A.*
Luminita Patron / *"I. G. Murgulescu" Institute of the Romanian Academy, Bucharest, Romania*
Derek Patton / *University of Houston, Houston, Texas, U.S.A.*
Sandrine Pensec / *Université Pierre et Marie Curie, Paris, France*
John H. Perepezko / *University of Wisconsin, Madison, Wisconsin, U.S.A.*
Myriam Perez De la Rosa / *The University of Texas, El Paso, Texas, U.S.A.*
Christopher C. Perry / *The Johns Hopkins University, Baltimore, Maryland, U.S.A.*
Claire S. Peyratout / *Max Planck Institute of Colloids and Interfaces, Potsdam, Germany*
M. P. Pileni / *Université Pierre et Marie Curie (Paris VI), Paris, France*
M. A. Pimenta / *Universidade Federal de Minas Gerais, Belo Horizonte, Brazil*
Frédéric Pincet / *Laboratoire de Physique Statistique de l'Ecole Normale Supérieure, Paris, France*
Laura Pirondini / *Università di Parma, Parma, Italy*
Bene Poelsema / *University of Twente, Enschede, The Netherlands*
Karsten Pohl / *University of New Hampshire, Durham, New Hampshire, U.S.A.*
Vojislava Pophristic / *University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*
Daniel Portinha / *Université Pierre et Marie Curie, Paris, France*
B. L. V. Prasad / *Kansas State University, Manhattan, Kansas, U.S.A.*

Maurizio Prato / *Università di Trieste, Trieste, Italy*
Aaron Puzder / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Georgios Pyrgiotakis / *University of Florida, Gainesville, Florida, U.S.A.*
Clayton J. Radke / *University of California, Berkeley, California, U.S.A.*
Srikanth Raghunathan / *Nanomat, Inc., North Huntingdon, Pennsylvania, U.S.A.*
Krishnaswami S. Raja / *The Scripps Research Institute, La Jolla, California, U.S.A.*
S. Ramakrishnan / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
G. V. Rama Rao / *The University of New Mexico, Albuquerque, New Mexico, U.S.A.*
David M. Rampulla / *Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.*
Lev Rapoport / *Holon Academic Institute of Technology, Holon, Israel*
Jean-Yves Raty / *University of Liège, Sart-Tilman, Belgium*
Serge Ravaine / *Centre de Recherche Paul Pascal, CNRS, Pessac, France*
Stéphane Reculusa / *Centre de Recherche Paul Pascal, CNRS, Pessac, France*
Mark A. Reed / *Yale University, New Haven, Connecticut, U.S.A.*
David N. Reinhoudt / *University of Twente, Enschede, The Netherlands*
Mikas Remeika / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
Maja Remskar / *Jozef Stefan Institute, Ljubljana, Slovenia*
Zhifeng Ren / *Boston College, Chestnut Hill, Massachusetts, U.S.A.*
Ryan M. Richards / *International University of Bremen, Bremen, Germany*
Carlos Rinaldi / *University of Puerto Rico, Mayagüez, Puerto Rico*
P. J. Roberts / *BlazePhotonics Ltd., Bath, United Kingdom*
Jose A. Rodriguez / *Brookhaven National Laboratory, Upton, New York, U.S.A.*
Miguel Rodriguez / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*
Andrey Rogachev / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
John A. Rogers / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*
Harry W. Rollins / *Idaho National Engineering and Environmental Laboratory, Idaho Falls, Idaho, U.S.A.*
Björn T. Rosner / *University of Wisconsin, Madison, Wisconsin, U.S.A.*
Vincent M. Rotello / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*
Wielsaw J. Roth / *ExxonMobil Research and Engineering Company, Annandale, New Jersey, U.S.A.*
Jason H. Rouse / *NASA-Langley Research Center, Hampton, Virginia, U.S.A.*
James F. Rusling / *University of Connecticut, Storrs, Connecticut, U.S.A.*
P. St.J. Russell / *University of Bath, Bath, United Kingdom*
Ajit Sadana / *University of Mississippi, University, Mississippi, U.S.A.*
R. Saito / *University of Electro-Communications, Tokyo, Japan*
Mark R. Sambrook / *University of Oxford, Oxford, United Kingdom*
Lynne A. Samuelson / *University of Massachusetts, Lowell, Massachusetts, U.S.A.*
Charlene A. Sanders / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*
Francesco Sansone / *Università di Parma, Parma, Italy*
Venugopal Santhanam / *Universität Duisburg-Essen, Essen, Germany*
Amitav Sanyal / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*
D. Y. Sasaki / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*
Daniel M. Schaadt / *University of California—San Diego, La Jolla, California, U.S.A.*
L. S. Schadler / *Rensselaer Polytechnic Institute, Troy, New York, U.S.A.*
David J. Schiffrin / *University of Liverpool, Liverpool, United Kingdom*
Jacob J. Schmidt / *University of California, Los Angeles, California, U.S.A.*
James W. Schneider / *Carnegie Mellon University, Pittsburgh, Pennsylvania, U.S.A.*

Holger Schönherr / *University of Twente, Enschede, The Netherlands*
Christian Schüller / *Universität Hamburg, Hamburg, Germany*
Purnesh Seegopaul / *Umicore USA Inc., Hillsborough, New Jersey, U.S.A.*
Jorge M. Seminario / *University of South Carolina, Columbia, South Carolina, U.S.A.*
Mamoru Senna / *Keio University, Yokohama, Japan*
Jonathan L. Sessler / *The University of Texas at Austin, Austin, Texas, U.S.A.*
Kay Severin / *École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
Sean E. Shaheen / *National Renewable Energy Laboratory (NREL), Golden, Colorado, U.S.A.*
Ehud Shapiro / *Weizmann Institute of Science, Rehovot, Israel*
Yury E. Shapiro / *Bar-Ilan University, Ramat-Gan, Israel*
P. C. Sharma / *University of California, Los Angeles, California, U.S.A.*
Pramod K. Sharma / *Nanomat, Inc., North Huntingdon, Pennsylvania, U.S.A.*
Roy Shenhar / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*
Peter M. A. Sherwood / *Kansas State University, Manhattan, Kansas, U.S.A.*
Masashi Shiraishi / *SONY Corporation, Yokohama, Japan*
Tatiana V. Shishkanova / *Institute of Chemical Technology, Czech Republic*
Ulrich Siggel / *Technischen Universität Berlin, Berlin, Germany*
Wolfgang M. Sigmund / *University of Florida, Gainesville, Florida, U.S.A.*
Richard Silberglitt / *RAND Corporation, Arlington, Virginia, U.S.A.*
Kevin Sill / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*
Amit Singhal / *NEI Corporation, Piscataway, New Jersey, U.S.A.*
Rhonda Skaggs / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Ganesh Skandan / *NEI Corporation, Piscataway, New Jersey, U.S.A.*
Gregory Ya. Slepyan / *Belarus State University, Minsk, Belarus*
Gabriela Smeureanu / *Hunter College of the City University of New York, New York, New York, U.S.A.*
James N. Smith / *National Center for Atmospheric Research, Boulder, Colorado, U.S.A.*
Karl Sohlberg / *Drexel University, Philadelphia, Pennsylvania, U.S.A.*
G. S. Solomon / *Stanford University, Stanford, California, U.S.A.*
Michael J. Solomon / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
Anongnat Somwangthanaroj / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
Wenbo Song / *Hokkaido University, Sapporo, Japan*
A. G. Souza Filho / *Universidade Federal do Ceará, Fortaleza, Brazil*
Patrick T. Spicer / *The Procter & Gamble Company, West Chester, Ohio, U.S.A.*
Ivan Stensgaard / *University of Aarhus, Aarhus, Denmark*
Geoffrey W. Stevens / *University of Melbourne, Melbourne, Victoria, Australia*
Rebecca L. Stoermer / *Pennsylvania State University, University Park, Pennsylvania, U.S.A.*
Savka I. Stoeva / *Kansas State University, Manhattan, Kansas, U.S.A.*
Peter K. Stoimenov / *Kansas State University, Manhattan, Kansas, U.S.A.*
Steven H. Strauss / *Colorado State University, Fort Collins, Colorado, U.S.A.*
Pieter Stroeve / *University of California, Davis, California, U.S.A.*
Makram T. Suidan / *University of Cincinnati, Cincinnati, Ohio, U.S.A.*
Gleb B. Sukhorukov / *Max-Planck-Institute of Colloids and Interfaces, Potsdam/Golm, Germany*
Bobby G. Sumpter / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*
Kai Sun / *University of Illinois at Chicago, Chicago, Illinois, U.S.A.*
Guangyu Sun / *National Cancer Institute, Frederick, Maryland, U.S.A.*
Xiangcheng Sun / *The University of Alabama, Tuscaloosa, Alabama, U.S.A.*

Aydin K. Sunol / *University of South Florida, Tampa, Florida, U.S.A.*
Sermin G. Sunol / *University of South Florida, Tampa, Florida, U.S.A.*
Robert E. Synovec / *University of Washington, Seattle, Washington, U.S.A.*
Makoto Tadokoro / *Osaka City University, Osaka, Japan*
Atsushi Takahara / *Kyushu University, Fukuoka, Japan*
Taishi Takenobu / *SONY Corporation, Yokohama, Japan*
Hideki Tanaka / *Okayama University, Okayama, Japan*
Zhiyong Tang / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
P. Tarakeshwar / *Pohang University of Science and Technology, Pohang, South Korea*
Jean-Marie Tarascon / *Université de Picardie—Jules Verne, Amiens, France*
Vandana R. Tarigopula / *University of South Carolina, Columbia, South Carolina, U.S.A.*
Juan M. D. Tascón / *Instituto Nacional del Carbón, CSIC, Oviedo, Spain*
Boyd R. Taylor / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Reshef Tenne / *Weizmann Institute, Rehovot, Israel*
Frederick Tepper / *Argonide Corporation, Sanford, Florida, U.S.A.*
Roger Terrill / *San Jose State University, San Jose, California, U.S.A.*
Mauricio Terrones / *Institute Potosino of Scientific and Technological Research (IPICYT), San Luis Potosí, Mexico*
V. A. Tertykh / *National Academy of Sciences of Ukraine, Kyiv, Ukraine*
Edwin L. Thomas / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*
Gregory B. Thompson / *The Ohio State University, Columbus, Ohio, U.S.A.*
Steven M. Thornberg / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*
Bernd Tieke / *Universität zu Köln, Köln, Germany*
Tia Benson Tolle / *Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio, U.S.A.*
Anna W. Topol / *IBM T. J. Watson Research Center, Yorktown Heights, New York, U.S.A.*
Jessica Torres / *The Johns Hopkins University, Baltimore, Maryland, U.S.A.*
Naoki Toshima / *Tokyo University of Science, Yamaguchi, Japan*
James M. Tour / *Rice University, Houston, Texas, U.S.A.*
Ali Toutianoush / *Universität zu Köln, Köln, Germany*
Günter E. M. Tovar / *University of Stuttgart, Stuttgart, Germany*
Costas Tsouris / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*
Yi Tu / *Boston College, Chestnut Hill, Massachusetts, U.S.A.*
Royale S. Underhill / *Defence Research and Development Canada—Atlantic, Dartmouth, Nova Scotia, Canada*
Rocco Ungaro / *Università di Parma, Parma, Italy*
Eberhard Unger / *Institute of Molecular Biotechnology, Jena, Germany*
Kohei Uosaki / *Hokkaido University, Sapporo, Japan*
Satoshi Utsunomiya / *University of Michigan, Ann Arbor, Michigan, U.S.A.*
Martin Valík / *Institute of Chemical Technology, Prague, Czech Republic*
Kevin E. Van Cott / *Virginia Polytechnic Institute and State University, Blacksburg, Virginia, U.S.A.*
G. Julius Vancso / *University of Twente, Enschede, The Netherlands*
Richard A. van Delden / *University of Groningen, Groningen, The Netherlands*
Daniel W. van der Weide / *University of Wisconsin, Madison, Wisconsin, U.S.A.*
Paul R. Van Tassel / *Yale University, New Haven, Connecticut, U.S.A.*
Natalia Varaksa / *University of Colorado, Boulder, Colorado, U.S.A.*
James C. Vartuli / *ExxonMobil Research and Engineering Company, Annandale, New Jersey, U.S.A.*
Stephanie Butler Velegol / *Pennsylvania State University, State College, Pennsylvania, U.S.A.*
Orlin D. Velev / *North Carolina State University, Raleigh, North Carolina, U.S.A.*

Alexei Yu. Vinogradov / *Osaka City University, Osaka, Japan*
Viriya Vithayaveroj / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
Viola Vogel / *University of Washington, Seattle, Washington, U.S.A.*
Pete Vukusic / *Exeter University, Exeter, United Kingdom*
Anthony J. Wagner / *The Johns Hopkins University, Baltimore, Maryland, U.S.A.*
Chien M. Wai / *University of Idaho, Moscow, Idaho, U.S.A.*
Joseph Wang / *New Mexico State University, Las Cruces, New Mexico, U.S.A.*
Kang L. Wang / *University of California, Los Angeles, California, U.S.A.*
Qian Wang / *University of South Carolina, Columbia, South Carolina, U.S.A.*
Shaopeng Wang / *Nomadics Inc., Stillwater, Oklahoma, U.S.A.*
Zhong Lin Wang / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
Paul Watts / *University of Hull, Hull, United Kingdom*
Achim Weber / *University of Stuttgart, Stuttgart, Germany*
Thomas J. Webster / *Purdue University, West Lafayette, Indiana, U.S.A.*
Brandon L. Weeks / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Alexander Wei / *Purdue University, West Lafayette, Indiana, U.S.A.*
Jie Wen / *The Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.*
Joachim H. Wendorff / *Philipps-Universität Marburg, Marburg, Germany*
Carter T. White / *Naval Research Laboratory, Washington, District of Columbia, U.S.A.*
George M. Whitesides / *Harvard University, Cambridge, Massachusetts, U.S.A.*
James Wicksted / *Oklahoma State University, Stillwater, Oklahoma, U.S.A.*
J. P. Wilcoxon / *Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.*
Andrew Williamson / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Marc Wirtz / *PPG Industries, Inc., Monroeville, Pennsylvania, U.S.A.*
Daniel B. Wolfe / *Harvard University, Cambridge, Massachusetts, U.S.A.*
Glenn M. Wolfe / *The Johns Hopkins University, Baltimore, Maryland, U.S.A.*
Stanislaus S. Wong / *State University of New York at Stony Brook, Stony Brook, New York, U.S.A.* / *Brookhaven National Laboratory, Upton, New York, U.S.A.*
Bruce W. Woods / *Lawrence Livermore National Laboratory, Livermore, California, U.S.A.*
Herbert Wormeester / *University of Twente, Enschede, The Netherlands*
Fanxin Wu / *University of California, Santa Cruz, California, U.S.A.*
H. Wu / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*
Jianzhong Wu / *University of California, Riverside, California, U.S.A.*
You-Ting Wu / *North Carolina State University, Raleigh, North Carolina, U.S.A.*
Chuanjun Xia / *University of Alabama at Birmingham, Birmingham, Alabama, U.S.A.*
Wei Xing / *The University of Queensland, Brisbane, Queensland, Australia*
Zhi Ping Xu / *University of North Texas, Denton, Texas, U.S.A.*
Dalia G. Yablon / *ExxonMobil Research and Engineering, Annandale, New Jersey, U.S.A.*
Boris I. Yakobson / *Rice University, Houston, Texas, U.S.A.*
Juchao Yan / *Eastern New Mexico University, Portales, New Mexico, U.S.A.*
Z. F. Yan / *University of Petroleum, Dongying, China*
Bai Yang / *Jilin University, Changchun, People's Republic of China*
Chao Yang / *Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.*
Jun Yang / *University of Alberta, Edmonton, Alberta, Canada*
Kun-Lin Yang / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*
T. Yang / *The Femtosecond Technology Research Association (FESTA), Ibaraki, Japan*

Wassana Yantasee / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*

Xiang-Rong Ye / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.* / *University of Idaho, Moscow, Idaho, U.S.A.*

Sotira Yiacoumi / *Georgia Institute of Technology, Atlanta, Georgia, U.S.A.*

Shiyoshi Yokoyama / *Communications Research Laboratory, Kobe, Japan*

Takashi Yokoyama / *Yokohama City University, Yokohama, Japan* / *National Institute for Materials Science, Tsukuba, Japan*

Seunghoo Yoo / *University of Massachusetts, Amherst, Massachusetts, U.S.A.*

Yi Yu / *Chinese Academy of Sciences, Changchun, People's Republic of China*

Markus Zahn / *Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.*

Mondona Zangeneh / *San Jose State University, San Jose, California, U.S.A.*

Stefan Zauscher / *Duke University, Durham, North Carolina, U.S.A.*

T. S. Zemanian / *Pacific Northwest National Laboratory, Richland, Washington, U.S.A.*

Hua Chun Zeng / *National University of Singapore, Singapore*

Dajie Zhang / *The Johns Hopkins University, Baltimore, Maryland, U.S.A.*

Gang Zhang / *Jilin University, Changchun, People's Republic of China*

Jin Z. Zhang / *University of California, Santa Cruz, California, U.S.A.*

Junhu Zhang / *Jilin University, Changchun, People's Republic of China*

Kai Zhang / *Jilin University, Changchun, People's Republic of China*

X. B. Zhang / *The University of Texas at Austin, Austin, Texas, U.S.A.*

Xunli Zhang / *University of Hull, Hull, United Kingdom*

Zongtao Zhang / *Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.*

Bin Zhao / *University of California, Riverside, California, U.S.A.*

Genady Zhavnerko / *The National Academy of Sciences of Belarus, Minsk, Belarus*

Xiange Zheng / *Drexel University, Philadelphia, Pennsylvania, U.S.A.*

N. B. Zhitenev / *Lucent Technologies, Murray Hill, New Jersey, U.S.A.*

Zhigang Zhou / *Duquesne University, Pittsburgh, Pennsylvania, U.S.A.*

Shu Zhu / *Princeton University, Princeton, New Jersey, U.S.A.*

C. F. Zukoski / *University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.*

Preface

Nanoscience encompasses all scientific phenomena that transpire in dimensions spanning the range of multiple atom clusters, molecular aggregates, supermolecular structures, polymers and biomolecules. In other words, nanoscience is the science of the nanoscale—or the very, very minute. Nanotechnology, the technological use of these properties and phenomena, has the potential to revolutionize a breathtaking range of fields, across practically all domains of human activity. The intense interest in using nanostructures stems from the promise that they boast superior electrical, optical, mechanical, chemical, or biochemical properties. There is little doubt that the use of these new materials and systems will lay the ground for the new technological revolution of the 21st century. Research in nanoscience is exploding as scientists from chemistry, physics, and biology, including colloid and polymer chemistry, materials and surface science, and biochemistry, biophysics and molecular biology have begun to examine the superior properties that characterize materials and phenomena at the nanoscale.

The *Dekker Encyclopedia of Nanoscience and Nanotechnology* is a vehicle by which scientists and the interested public can explore the most recent developments in today's nanomaterials, and preview several of their foreseen applications of tomorrow. The subject areas of this five-volume collective work include, but are not limited to, chemistry, physics, molecular and quantum biology, materials science and engineering, and computer science. The topics include fullerenes and carbon nanotubes, quantum dots and inorganic nanowires, interfacial chemistry, nanostructures, analytical and characterization methods, design and fabrication techniques of nanocomposites, properties of functional nanomaterials, nanostructured catalysts, molecular electronics, optical devices, bionanotechnology, colloidal phenomena—even future and environmental considerations about nanotechnology. In short, the reference strives to encompass, document, and organize the enormous proliferation of information that has emerged from the revolution at the nanoscale.

It is the editors' hopes that the *Dekker Encyclopedia of Nanoscience and Nanotechnology* will help researchers, graduate students, undergraduate students, and nonprofessionals all better understand and participate in this remarkable emerging field. To keep the reference accessible, the *Encyclopedia* is published in both online and print formats. The print version consists of multiple traditional hardbound volumes with articles arranged alphabetically. The front matter to each volume provides an alphabetical Table of Contents and a Topical Table of Contents. An index at the back of each volume is intended as a further guide.

The online version of the *Encyclopedia* has been created by coupling the content of the print edition with a powerful search engine and user-friendly interface. The online database is dynamic, with additional articles added each quarter. As with the print edition, users will be able to browse the alphabetical and topical Table of Contents, but they will also be able to search for entries by keywords.

The editors of the *Dekker Encyclopedia of Nanoscience and Nanotechnology* could not have accomplished their feat without the help of many. We first thank the authors of more than 300 articles which, as recognized experts in their fields, lend their credibility and prestige to the *Encyclopedia*. In addition, the editors were joined by an esteemed International Editorial Advisory Board whose input was crucial to shaping the reference. Success in the coordination of the activities that has resulted in this final product is due to the remarkable staff at Marcel

Dekker, Inc. We are indebted to the direction provided to us by Ellen Lichtenstein, Carolyn Hall, Oona Schmid, Susan Lee, Joanne Jay, and Jeeran Ok. In particular, the creative ideas and boundless energy that Oona Schmid has demonstrated is greatly acknowledged by the three of us. We also thank Anita Lekhwani for her confidence in us as editors and Russell Dekker for his support and encouragement.

James A. Schwarz
Syracuse University, Syracuse, New York

Cristian I. Contescu
Material Methods LLC, Newport Beach, California

Karol Putyera
Shiva Technologies, Syracuse, New York

A Note on Terminology

New terms, based on the prefix ***nano***, have gained broad circulation, unified the older terminology, and defined the topics of newly found disciplines. Just as the research community has struggled to restrict when and where the prefix ***nano*** can be appropriately used, we too struggled with the *Dekker Encyclopedia of Nanoscience and Nanotechnology* to avoid puncturing every sentence with the prefix ***nano***. As this terminology develops, the online edition of the *Encyclopedia* will be able to incorporate these new conventions.

But at the time of publication of the first edition of the *Encyclopedia* these questions are still unresolved. One definition that has been proposed in order to remove the ambiguity would limit the use of the ***nano*** prefix to research and development of objects having the key property that they have at least one dimension in the range of 1 to 100 nanometers. New advances in synthetic methods for making such structures, combined with new analytical and manipulation tools, made it possible to refine methods to the point where *de novo* designed objects with nanoscopic dimension can, in many cases, be assembled in molecule-pure form or spatially organized structures. These systems, designed through processes that exhibit fundamental control over the physical and chemical attributes of molecular-scale structures, can be combined to form larger structures.

However, the *Dekker Encyclopedia of Nanoscience and Nanotechnology* has numerous entries that include the words micro and meso. Thus the use of the term ***nano***, which, according to one broadly circulated definition, only limits research and technology to development in the length scale of approximately 1–100 nanometer range, is not simply a metric of length. Science at the nanoscale does not accept rigid limits on dimensionality. Indeed the very utility of nanoscience may be compromised by arbitrarily circumscribing its reach. After all, for ancient Greeks, the term "nanos" meant a dwarf. Keeping this in mind, we attempted in this reference to use the term ***nano*** to refer to objects and their subsequent processing into newly created structures, devices, or systems that have novel properties and functions because of their small and/or intermediate size. In other words, **size and performance** are the critical parameters that provide the requisite conditions to justify the use of the term ***nano***.

By adopting a more elastic definition, which on one hand spans sizes from a few nanometer(s) to one (or a few) hundred(s) nanometers, but at the same time recognizes that the properties and performance of nanoobjects are rooted in their nanoscopic size, we believe that we made justice to all views that currently shape this field of continuous development and hope that other investigators—at universities, state laboratories, and in industries—will follow our lead.

Contents

Volume 2

Volume 3

Volume 4

Volume 5

Topical Contents

Carbon Nanotubes

Catalysts and Catalysis at the Nanoscale

Colloidal Phenomena

Computational and Theoretical Modeling

Design and Fabrication of Nanocomposites

Environmental Applications and Societal Implications

Fullerenes

Functional Nanomaterials

Inorganic Nanowires

Molecular Electronics and Devices

Nanostructures and Nanophase Structures

Optical Devices

Quantum Dots

Self Assembly and Self Organization

Sensors and Actuators

Supramolecular Chemistry

Molecular Designs for Self-Organized Superstructures

Surface and Interfacial Chemistry

Near-Field Microscopy Techniques

Björn T. Rosner
Daniel W. van der Weide
University of Wisconsin, Madison, Wisconsin, U.S.A.

INTRODUCTION

While force interactions are most commonly employed for probing, manipulation, and addressing of nanoscale features, dielectric contrast mechanisms are also potentially useful means of probing nanostructures. Optical absorption or emission, dielectric contrast in the infrared (IR), far-infrared (FIR), or microwave regime are all part of a rich pool of phenomena observable with electromagnetic interactions, with applications in nanoelectronics, spintronics, and biotechnology being developed. To attain the highest resolution, it is necessary to confine the area in which the probing field interacts with the sample to subwavelength dimensions. This is accomplished by fabricating probes with subwavelength-sized features and scanning such near-field probes very close to the sample. This article summarizes such near-field techniques in the visible, IR, and microwave regime and relates them to current and future applications.

HISTORICAL OVERVIEW

Doing microscopy with electromagnetic waves usually implies using visible light. This is because Abbe's[1] much-cited criterion limits resolving power to $\sim\lambda/2$ (where λ is the wavelength), so using millimeter wave and infrared frequencies would result in unattractively low resolution. While this criterion is correct for far-field microscopy, it is not limiting when near-field interactions are taken into account. Theoretical considerations were first published by Synge[2] in 1928 and Bethe[3] in 1944. The first practical implementation of near-field microscopy was developed for radio frequency applications by Frait[4] (1959) and Soohoo[5] (1962). Further important milestones were the first use of a coaxial waveguiding probe to localize microwave fields for measuring material properties by Bryant and Gunn[6] in 1965 and $\lambda/60$ wavelength-relative resolution achieved by Ash and Nichols[7] in 1972 using a subwavelength aperture. The advent of scanning tunneling microscopy (STM)[8] and subsequent scanning probe techniques provided a means by which accurate nanoscale tip–sample distance control became possible for the first optical near-field microscopes.[9,10] Finally, the advent of scanning force microscopy (SFM) in 1986[11] allowed for near-field scanning optical microscopy (NSOM or SNOM) instruments relying on (shear-) force feedback[12] in 1992. From that point, near-field microscopy has spread through a wealth of techniques and applications to form a diverse and very active field of research.

SCANNING NEAR-FIELD OPTICAL MICROSCOPY

Fiber Probes

For visible wavelengths, the most common probe type today uses tapered optical fibers that are coated with an opaque metal layer except at the apex where a small opening defines the optical output (Fig. 1). The shear-force feedback signal is commonly electronically detected using tuning fork feedback.[14] In comparison to SFM probes, straight SNOM fibers are much stiffer and more fragile, leading to slower scan speeds and difficulties in establishing reliable feedback, particularly on soft samples. Another major concern for optical near-field microscopy is the low-power throughput of NSOM fibers because they act as cutoff waveguides for light near the apex. To improve transmission over that of the common technique of pulling heated fibers, meniscus etching[15] and double-tapered probes have been used.[16] Tube etching, a technique where the polymer coating of the fiber is not removed prior to etching, has been introduced.[17] Typical applications for fiber-NSOM instruments are found in chemistry, biology, material science, and integrated optics, e.g., single molecule spectroscopy,[18] cell membrane structure[19] analysis, and optical waveguide mapping[20] (Fig. 2). Another field of activity is to use NSOM as a means for high-density optical data storage.[21] Further information regarding aperture NSOM can be found in two review articles by Hecht et al.[22] and Dunn.[23]

Apertureless Probes

Apertureless SNOM is the complement to aperture SNOM. Apertureless SNOM includes a wide variety of

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009295

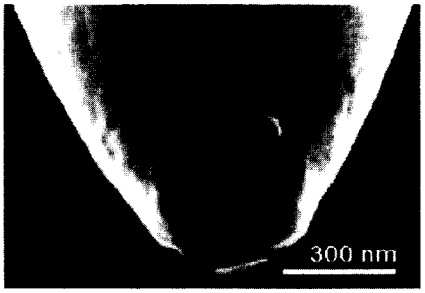


Fig. 1 A scanning electron micrograph of a typical tapered aluminum-coated optical fiber tip. (From Ref. [13]. Copyright 1995 American Institute of Physics.)

related approaches, including the use of surface plasmons and scattering-type SNOM (s-SNOM). While the use of plasmon-guiding effects to localize light is still in its infancy, considerable interest in this field exists[24] because light throughput might be increased by orders of magnitude, e.g., through the use of a tetrahedral tip with an integrated surface-plasmon guide as a scanning probe.[25]

The s-SNOM technique is a prime example of apertureless SNOM.[26] Here a sharp probe, typically an STM or SFM probe, is scanned over an illuminated surface (Fig. 3). The optical contrast of s-SNOM arises from the

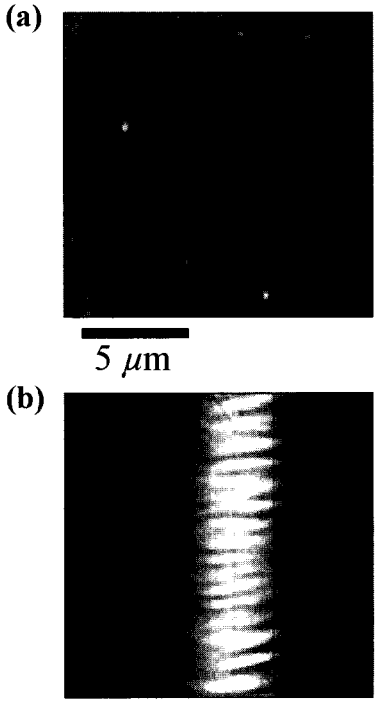


Fig. 2 Topography (a) and optical intensity map (b) of the TM mode in a Ti:$LiNbO_3$ waveguide, taken using a typical fiber NSOM setup. (From Ref. [20]. Copyright 2002 American Institute of Physics.)

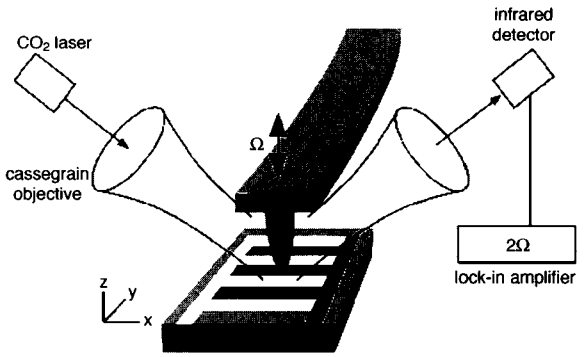


Fig. 3 Sketch of a scattering-type near-field optical microscope using higher-order modulation for background suppression. (From Ref. [27]. Copyright 2000 American Institute of Physics.)

dipole–dipole interaction between the sharp tip and the sample, leading to characteristic local scattering [more about the theory can be found in the section about infrared (IR) near-field microscopy]. The main advantage of s-SNOM is that the resolution is not limited by an aperture or the skin depth of metal, but rather by the sharpness of the probing tip, potentially leading to better resolution. This advantage has been used to localize polymerization of commercial photoresists in an effort to push the resolution of optical lithography. Resolutions down to 70 nm were recently reported using this technique.[28] An s-SNOM setup was also used to simultaneously record optical amplitude and phase contrast on a subwavelength scale.[29]

Integrated Probes

Finally, researchers have been microfabricating SNOM probes that could be applied in the same fashion as cantilevered SFM probes, with minor instrumental changes. Some of the proposed concepts are passive, such as metal-coated probes with circular[30] apertures and solid immersion lenses on cantilevers[31] (strictly speaking a far-field technique). Initial efforts have been made to fabricate active devices in GaAs that include integrated light sources.[32] Considerably more research has been performed to produce photodetectors integrated close to the probe[33,34] or right at the apex of the sharp tip,[35] preferably integrated with the cantilever and chip body.[36]

SCANNING NEAR-FIELD INFRARED AND FAR-INFRARED MICROSCOPY

Because of the wealth of chemical information available through vibrational band spectroscopy, many traditional infrared techniques such as Fourier-transform infrared spectroscopy (FTIR), thermal probing, and Raman scat-

tering have become cornerstones of chemical characterization. In recent years, characterization of optoelectronic devices in the IR also gained importance for probing integrated-circuit lasers. Proximal probes offer a way to combine IR imaging with subwavelength resolution. Fiber-based aperture scanning near-field infrared microscopy (SNIM) setups, similar to optical fiber NSOM instruments, have been developed, based on special infrared transmissive fibers and etching techniques.[37]

The far-infrared (FIR) spectrum (10 μm$<\lambda<$1 mm) is of particular interest because it is part of the molecular fingerprint regime. Here characteristic molecular rotational or vibrational absorption spectra offer the promise of chemical microscopy. Biological membrane absorption and conductive or dielectric properties of materials, e.g., superconductors and quantum dots, offer additional applications for FIR near-field microscopy. Little work has been performed in this frequency regime, mostly because high-power FIR sources and sensitive FIR detection are not yet readily available.

Aperture Techniques

One practical FIR technique is to use a three-dimensional funnel or tapered cylindrical waveguide to confine radiation in an aperture approach. This geometry was employed[38] to overcome the imperfect focusing capabilities of traditional FIR optics. Power transmittance in the cutoff region is

$$T = e^{-4\pi\sqrt{(f_c^2-f^2)}t}$$

where t is the penetration depth into the cutoff region, and f and f_c are the frequency of the radiation and the cutoff frequency, respectively. In practice, a sharp cutoff in transmitted intensity was found at an aperture size of about $\lambda/2$. This setup was then applied for far-infrared near-field spectroscopy of two-dimensional electron systems.[38] Aperture confinement of FIR radiation has been one of the leading techniques to date[39,40] because of its straightforward implementation. Keilmann et al. showed that by introducing a center conductor into the tapered hollow guide and creating an in situ coaxial waveguide, the aperture size can be decreased by almost an order of magnitude without losing transmitted power (Fig. 4). It was also found that the transmitted power varies with a period of $\lambda/2$ as the length of wire protruding from the aperture is extended because it acts as a resonant antenna.

Apertureless Techniques

Using apertureless (scattering) SNIM (Fig. 3) in the IR and FIR has even better prospects than in the visible regime. For one, because the resolution depends on the scattering tip, not the wavelength, the wavelength-specific

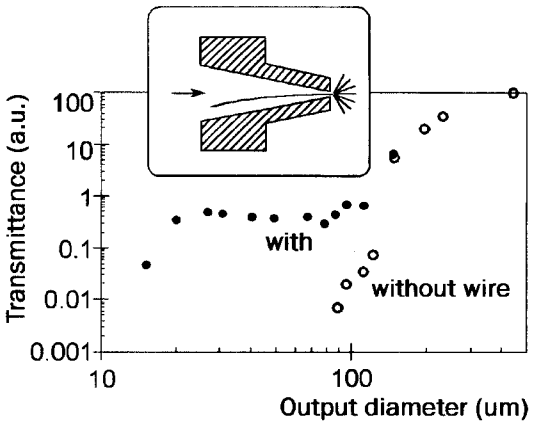


Fig. 4 The relative transmittance at a wavelength of $\lambda=392$ μm of focusing cones with different output apertures is investigated. Adding a central, coaxial wire prevents cutoff reflection at $d<\lambda/2$. (From Ref. [39]. Copyright 1995 Elsevier Science.)

resolution is far better. Far-infrared microscopy with nanometer resolution is possible and has been demonstrated.[41,42] Furthermore, because the wavelength is much longer than the occasional distance variations between tip and sample, optical interference effects play less of a role, reducing the risk of topographical artifacts in the optical data. In scattering-type instruments, it is also possible to eliminate topography contrast and record pure optical contrast.[42,43] Here when applying noncontact SFM, the nonlinear dipole–dipole interaction between the tip and sample leads to (purely optical) higher modulation harmonics in the scattered signal, which can be used for lock-in detection. Following one argument,[42] image contrast arises from the interaction between the polarizable probe (modeled as a sphere with polarizability α and radius a) and the dielectric or metallic sample with complex dielectric number ε. This results in an effective polarizability of the coupled system

$$\alpha_{\perp}^{\text{eff}} = \frac{\alpha(1+\beta)}{1-\dfrac{\alpha\beta}{16\pi(z+a)^3}}$$

where $\beta=(\varepsilon-1)/(\varepsilon+1)$. This formula shows both strong enhancement at short probe–sample separation z and the existence of amplitude and phase effects because α and β are complex. Thus the scattered far field $E_{\text{sca}}=\alpha^{\text{eff}}E$ reports the phase and amplitude of the complex near-field interaction. With this technique, Knoll and Keilmann[27] have mapped conductivity of a silicon sample with nanometer resolution using a scattering-type near-field microscope and 10.6-μm radiation as shown in Fig. 5. The same researchers[44] reported contrast between polystyrene (PS) and polymethylmethacrylate (PMMA) on the nanometer scale using the same instrument and two wavelengths

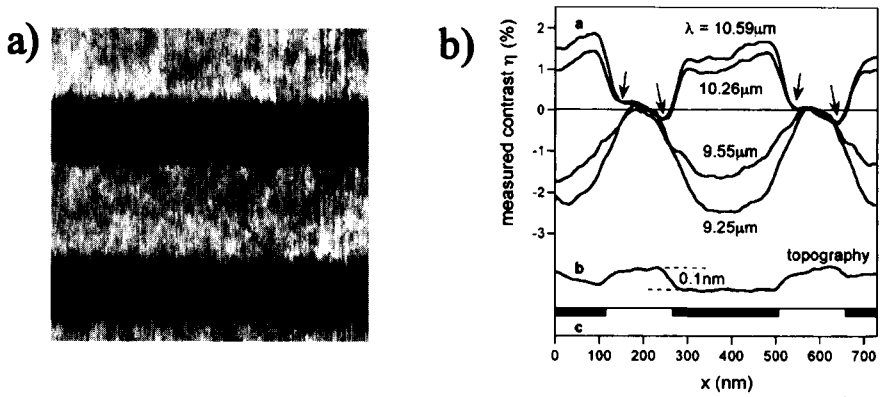


Fig. 5 (a) Infrared s-SNOM image of a flat Si sample with subsurface doping (bright), taken at $\lambda = 10.59$ μm. (b) Line scans taken across doping stripes, measured at four different wavelengths, demonstrating 30-nm spatial resolution and a contrast reversal when the infrared wavelength tunes through the induced plasmon resonance, here at about $\lambda = 10$ μm. (From Ref. [27]. Copyright 2000 American Institute of Physics.)

around 10 μm (Fig. 6). Hillenbrand et al.[45] applied this instrument to probe the local interaction with an SiC sample and found phonon-enhanced near-field coupling that is extremely sensitive to the chemical and structural composition of polar materials.

Additional Techniques

Another promising approach is to use antenna structures as near-field sources.[46,47] While any coaxial open-ended waveguide can be considered an antenna, guiding radiation at visible or infrared wavelengths is not yet technologically feasible. However, by using antennas that collect far-field radiation and channel it to near-field tips for exciting the sample, antenna concepts can be very effective. Other approaches for FIR near-field microscopy include THz time domain spectroscopy (TDS) using photoconductive THz emitters[40] pumped by femtosecond optical lasers and all-electronic generation of THz radiation with nonlinear transmission lines (NLTL).[48]

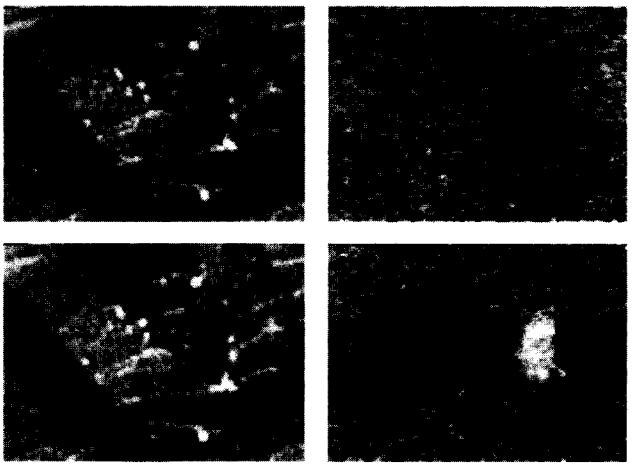


Fig. 6 Simultaneously acquired topography (left) and infrared images (right) of PS embedded in PMMA using an s-SNOM. The upper row is recorded with an illumination wavelength of $\lambda = 9.68$ μm, where PS has an absorption maximum. The lower row shows the same spot, illuminated at the PMMA absorption maximum wavelength of $\lambda = 10.17$ μm. The field of view is 3.5×2.5 μm. (From Ref. [44]. Copyright 1999 Nature.)

Aperture probes, scattering probes, and coaxial tips form the main body of research on near-field microscopy as it is applied in NSOM. An entirely different approach to infrared microscopy lies in the use of photothermal signals for near-field spectroscopy.

Here a thermal probe in temperature-sensing mode directly measures absorption of incident IR radiation by sensing the induced heating of the sample. This heating can either be directly measured for each wavelength in the dispersive approach[49] or a Fourier-transform absorption spectrum can be acquired when using intensity-modulated illumination.[50] Many other thermal techniques exist, e.g., scanning thermal microscopy (SThM) where the tip is used to heat the sample and sense heat flow out of the probe. More about localized thermal imaging can be found in a review by Kölzer et al.[51] Reviews of near-field infrared microscopy and near-field vibrational spectroscopy have been written by Dragnea and Leone[52] and by Pollock and Smith.[53]

SCANNING NEAR-FIELD MICROWAVE MICROSCOPY

Unlike in the visible or infrared regime, metals are good conductors for microwave frequencies. Metal wave-

guides are an efficient means of transmitting microwave power and can be subwavelength sized in cross section. Microwave near-field probes are either broadband or resonant, which can make them more sensitive, but also more limited in the frequency range in which they can be operated.

Coaxial Waveguides

An advantageous waveguide design is to use tapered coaxial waveguide probes that increase resolution and provide high throughput. For the highest resolution, a sharpened center helps concentrate the fringing fields emanating from the probe apex (lightning rod effect) and thus confines the interaction area between the probe and sample. Proof-of-principle experiments with a sample between two such sharpened coaxial transmission line tips by Keilmann et al.[54] resulted in $\lambda/10^6$ wavelength-relative resolution. Integrating coaxial structures with cantilevered SFMs enables microscale probing of passive or active samples in reflection or transmission mode, respectively. Initial probes by van der Weide et al.[55] were made by coating a commercial silicon SFM probe with photoresist and gold, then opening the coaxial structure by gently rubbing the tip on a substrate while in feedback. These probes allowed for local measurements of waveforms on fast integrated circuits. Microfabricated probes (Fig. 7)[36] have various advantages, among them higher resolution and batch fabrication. Probes made from GaAs report superior signal transfer.[57]

Scanning Tunneling Microscopy-Based Microwave Microscopy

Just as with SFM, STM can offer a tip–sample distance feedback system for microwave near-field microscopy while providing a sharp conducting tip to localize high-frequency electromagnetic interaction. One such system was developed by Kramer et al.[58] Here higher harmonic signals are generated because of tunnel junction nonlinearities. Because the higher-harmonic generation strongly depends on the electronic properties of the surface, it is possible to differentiate different chemical species. McCarty and Weiss[59] and Michel et al.[60] have explored this possibility in a number of publications. These instruments have been applied from d.c. to 20 GHz, recording linear or nonlinear (higher-harmonic)[61] microwave spectra. Usually, topography has been controlled using conventional STM while simultaneously recording millimeter wave amplitude in transmission or reflection mode.

Apertureless Approaches

Scattering of incident far-field microwave radiation from a conducting tip in close proximity to a sample is analogous to s-SNOM in the visible, IR, or FIR regions. Thus numerical and theoretical treatments of s-SNOM can be applied to the microwave regime if certain material constants, such as indices of refraction, are revised. Surprisingly, although the resolution of such a microscope is defined by the sharpness of the tip and can lie in the nanometer range, little research has been carried out in this field. Keilmann et al.[62] described the coaxial setup specified above[58] as essentially equivalent to an s-SNOM setup for the microwave regime. Instead of using far-field optics to focus the radiation, they use coaxial cables to direct the fields, which is much more convenient for microwave frequencies. They point out that the resolution achieved is already an order of magnitude higher than the microwave penetration depth of the tip material, indicating that no intrinsic physical mechanism seems to exist that would limit the resolution achievable.

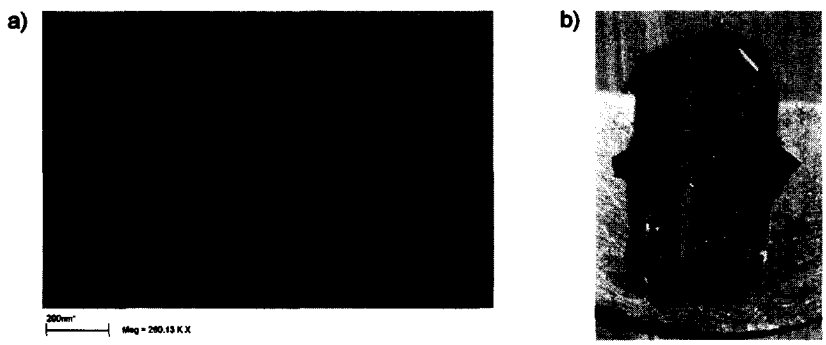


Fig. 7 (a) Scanning electron micrograph of a nanometer-scale Schottky diode at the apex of an SFM tip. (b) This SNOM detector is integrated on a cantilever and chip body that contains waveguides for straightforward connection to macroscopic instrumentation. (From Ref. [56]. Copyright 2002 American Institute of Physics.)

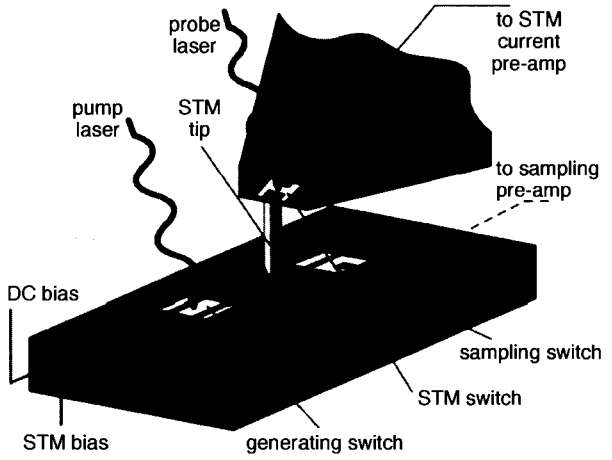


Fig. 8 Localized pump and probe setup that uses photoconductive switches to trigger a pulse on the DUT and to probe the instantaneous voltage on the DUT through the tip. (From Ref. [65]. Copyright 1996 American Institute of Physics.)

Ultrafast Scanning Tunneling Microscopy

Ultrafast scanning tunneling microscopy (USTM) was independently invented by Weiss et al.[63] and Nunes and Freeman[64] in 1993. Both techniques rely on pump and probe experiments for highest temporal resolution. Using femtosecond lasers that repetitively illuminate photoconductive (PC) switches ensures high temporal resolution, while using a sharp tip gives the additional benefit of high spatial resolution. The nonlinear current vs. voltage tunneling characteristics are employed to measure the picosecond cross-correlation. In the setup of Nunes, pump and probe travel along one transmission line while Weiss employs the more common setup of pumping the sample and probing the tunneling tip as shown in Fig. 8.

A related pump-probe technique was introduced by Kim et al.[66] that relies on direct conduction through a tip. They brought a probe with Ti tip, PC switch, and electric lead into contact with a transmission line. The convolution between pump pulse and probe pulse could be mapped out by probing the instantaneous electric field at the tip position by opening the PC switch and collecting the signal transmitted through the tip. Takeuchi and Mizuhara[67] improved on this concept by using an LT-GaAs tip and employing SFM feedback. A propagating pulse on a metal strip was mapped in 2-D and observed in sequential images with 0.8-psec time intervals.

Oesterschulze et al.[68] have also reported on an SFM setup with Lt-GaAs probe and PC switches. Here contact and noncontact images of a voltage pulse on a coplanar stripline were presented with picosecond temporal resolution.

Resonant Probes

Resonant structures are superior in that they allow localization of high field areas and are often more sensitive. However, most resonant techniques do not allow for resolution on the single micrometer scale or below. Among those larger-scale probes are narrow resonant slots in a rectangular hollow waveguide[69,70] and open resonant coaxial resonators.[71]

For higher resolution, Tabib-Azar et al.[72] have used a tapered stripline resonator for near-field microwave microscopy in a series of publications. The resonator is designed by capacitively coupling to the stripline, and a sharpened stainless steel wire was used as a probe tip. Because the strong distance dependence of the reflected signal is also present in this resonant setup, an independent distance control was employed. The researchers used a reflectance-compensated fiber optic distance sensor next to the tip and maintained the probe–sample separation with an accuracy of 0.4 μm. The researchers claim 0.4 μm spatial resolution at 1 GHz for this instrument.

High-resolution microwave near-field probes incorporating $\lambda/4$ coaxial resonators have been produced by the group of Xiang[73] using an STM-like tip as a point-like field emitter, protruding from a small aperture (Fig. 9). The sharp tip results from tapering the center conductor of the coaxial cavity. The instrument achieves a spatial

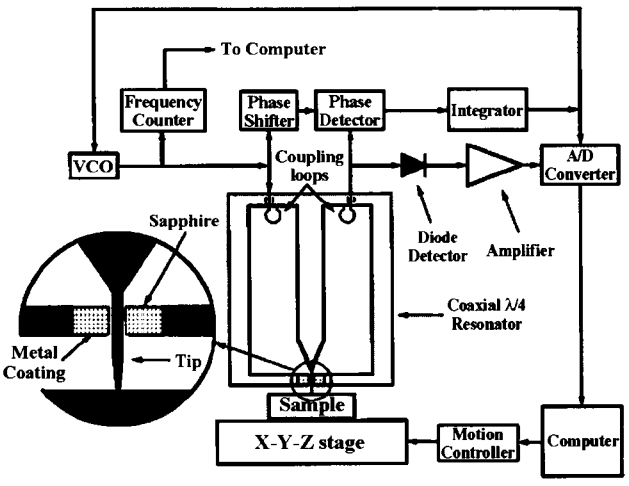


Fig. 9 Resonant SNMM setup using a coaxial cavity. (From Ref. [73]. Copyright 1997 American Institute of Physics.)

resolution of 100 nm and was used to image dielectric-constant profiles and ferroelectric domains.[74] Ferroelectric domains can be imaged using microwave fields by taking advantage of the nonlinear components in the dielectric constant for these materials. The contrast mechanism is based on a low-frequency oscillating voltage between sample and substrate that modulates the effective dielectric constant of the sample. The slight shift of the resonance frequency in the coaxial resonator probe could be related to this variation in dielectric constant, allowing for quantitative mapping of the third-order dielectric constant in a ferroelectric domain structure as seen in Fig. 10. To decrease tip damage resulting from excessive forces on the tip, and to be able to image unprepared conducting samples, Duewer et al.[75] introduced a tip–sample feedback scheme based on regulating the resonance frequency of the cavity to maintain a constant separation. In this experiment, conductivity was simultaneously measured via the amplitude of the cavity resonance. The principle for resistivity microscopy is based on the fact that the reflection of electromagnetic waves from conducting surfaces is determined by their resistivity. Therefore measuring the reflected amplitude and phase of a near-field probe yields a resistivity map.

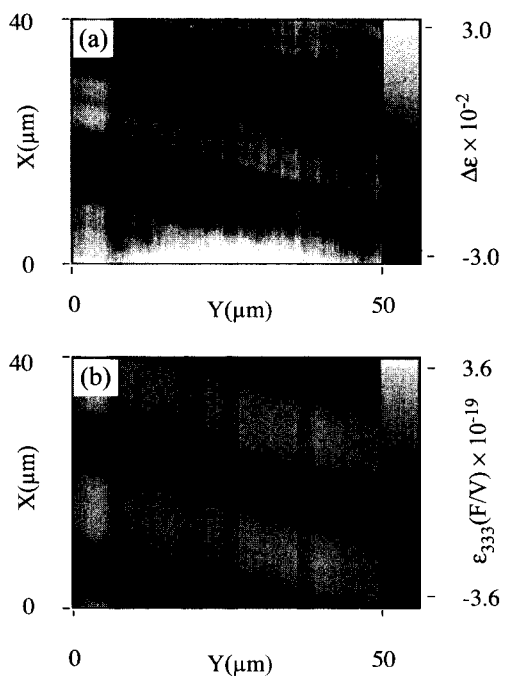


Fig. 10 Linear dielectric constant image (a) and nonlinear third-order dielectric constant image (b) of the same area of a $LiNbO_3$ single crystal with periodically polarized ferroelectric domains, recorded using a coaxial resonator SNMM. (From Ref. [74]. Copyright 1998 American Institute of Physics.)

The reflection coefficient of a plane wave from a conducting surface is

$$\Gamma = \left(\frac{Z_s - Z_0}{Z_s + Z_0}\right)$$

where Z_0 is the characteristic impedance of the line and Z_s is the surface impedance, which is directly connected to the conductivity:

$$Z_s = \sqrt{\frac{i\omega\mu_0}{\sigma + i\omega\varepsilon}}$$

where ω is the radial frequency, μ_0 is the permeability of free space, and ε is the dielectric constant. For thick conducting layers, much thicker than the skin depth, surface impedance is measured while thin layers yield sheet resistance. Another important application of evanescent microwave probing is to measure spatially resolved surface impedance of high-temperature superconductor films. By finding superconducting and normally conducting features on superconducting thin films, it is possible to localize defects and evaluate film quality. The contrast mechanism is based on variations in microwave absorption when keeping the sample close to the critical temperature.[76] Others employ room temperature normal-state measurements[77] and relate them to the film's superconducting properties.

Tabib-Azar et al.[78] have used evanescent microwave probes for several nontraditional probing applications. These include transient and steady state thermography on semiconducting materials with 0.01 K thermal and 1 μs temporal resolution. We recently published a review of near-field microscopy in the microwave and infrared regime.[56]

CONCLUSION

We have summarized recent and current trends in optical, infrared, and microwave near-field microscopy incorporating surface force feedback. In the visible regime, the most common techniques are fiber-based aperture microscopes and scattering-type setups using AFM tips. On the other side of the spectrum in the microwave regime, metallic waveguide geometries play a dominant role. Future developments will enable combined probing of topography and field at the nanometer scale. The development of tools and instruments for mechanical manipulation of structures at nanometer length scales has driven progress in nanoscale science and technology. At these scales, the doors to new worlds in biology, chemistry, and physics begin to open, enabling new applications such as real-time protein structure identification, molecular electronics, and

quantum computation. However, to approach these ambitious applications, a great instrumental challenge must be surmounted: Tools must not only be able to touch, hold, and move molecular-scale objects, but must also be able to identify and control molecules and molecular-scale objects such as nanowires based on their chemical species or function. Near-field probes that combine access to dielectric contrast mechanisms at the nanometer scale hold out this promise.

REFERENCES

1. Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Arch. Mikrosk. Anat. Entwickl.mech. **1873**, *9*, 413–468.
2. Synge, E.H. A suggested method for extending microscopic resolution into the ultra-microscopic region. Philos. Mag. **1928**, *6*, 356–362.
3. Bethe, H.A. Theory of diffraction by small holes. Phys. Rev. **1944**, *66* (7–8), 163–182.
4. Frait, Z. The use of high-frequency modulation in studying ferromagnetic resonances. Czechoslov. J. Phys. **1959**, *9*, 403–404.
5. Soohoo, R.F. A microwave magnetic microscope. J. Appl. Phys. **1962**, *33* (3), 1276–1277.
6. Bryant, C.A.; Gunn, J.B. Noncontact technique for the measurement of semiconductor resistivity. Rev. Sci. Instrum. **1965**, *36* (11), 1614–1617.
7. Ash, E.A.; Nicholls, G. Super-resolution aperture scanning microscope. Nature **1972**, *237* (5357), 510–512.
8. Binnig, G.; Rohrer, H. Scanning tunneling microscopy. Helv. Phys. Acta **1982**, *55* (6), 726–735.
9. Pohl, D.; Denk, W.; Lanz, M. Optical stethoscopy: Image recording with resolution $\lambda/20$. Appl. Phys. Lett. **1984**, *44* (7), 651–653.
10. Betzig, E.; Lewis, A.; Harootunian, A.; Isaacson, M.; Kratschmer, E. Near-field scanning optical microscopy (NSOM). Biophys. J. **1986**, *49*, 269.
11. Binnig, G.; Quate, C.F.; Gerber, C. Atomic force microscope. Phys. Rev. Lett. **1986**, *56* (9), 930–933.
12. Toledo-Crow, R.; Yang, P.C.; Chen, Y.; Vaez-Iravani, M. Near-field differential scanning optical microscope with atomic force regulation. Appl. Phys. Lett. **1992**, *60* (24), 2957–2959.
13. Obermüller, C.; Karrai, K. Far field characterization of diffracting circular apertures. Appl. Phys. Lett. **1995**, *67* (23), 3408–3410.
14. Karrai, K.; Grober, R.D. Piezo-electric tuning fork tip–sample distance control for near field optical microscopes. Ultramicroscopy **1995**, *61*, 197–205.
15. Hoffmann, P.; Dutoit, B.; Salathe, R.P. Comparison of mechanically drawn and protection layer chemically etched optical fiber tips. Ultramicroscopy **1995**, *61* (1–4), 165–170.
16. Saiki, T.; Mononobe, S.; Ohtsu, M.; Saito, N.; Kusano, J. Tailoring a high-transmission fiber probe for photon scanning tunneling microscope. Appl. Phys. Lett. **1996**, *68* (19), 2612–2614.
17. Stöckle, R.M.; Fokas, C.; Deckert, V.; Zenobi, R.; Sick, B.; Hecht, B.; Wild, U.P. High-quality near-field optical probes by tube etching. Appl. Phys. Lett. **1999**, *75* (2), 160–162.
18. Xie, S.; Lu, P.; Bian, R.; Sanchez, E.; Leung, P. In *Near-Field Single Molecule Spectroscopy*, Quantum Electronics and Laser Science Conference, Anaheim, CA, 1996; 256–257.
19. Hulst, N.F.v.; Moers, M.H.P. Biological applications of near-field optical microscopy. IEEE Eng. Med. Biol. **1996**, *15* (1), 51–58.
20. Campillo, A.L.; Hsu, J.W.P.; White, C.A.; Jones, C.D.W. Direct measurement of the guided modes in $LiNbO_3$ waveguides. Appl. Phys. Lett. **2002**, *80* (13), 2239–2241.
21. Minh, P.N.; Ono, T.; Tanaka, S.; Goto, K.; Esashi, M. Near-field recording with high optical throughput aperture array. Sens. Actuators, A **2002**, *95* (2–3), 168–174.
22. Hecht, B.; Sick, B.; Wild, U.P.; Deckert, V.; Zenobi, R.; Martin, O.J.F.; Pohl, D. Scanning near-field optical microscopy with aperture probes: Fundamentals and applications. J. Chem. Phys. **2000**, *112* (18), 7761–7774.
23. Dunn, R.C. Near-field scanning optical microscopy. Chem. Rev. **1999**, *99*, 2891–2927.
24. Fischer, U.C. Microsopic Electromagnetic Radiation Transmitter or Detector, US Patent #5,770,855, 1998.
25. Heimel, J.; Fischer, U.C.; Fuchs, H. SNOM/STM using a tetrahedral tip and a sensitive current-to-voltage converter. J. Microsc. **2001**, *202* (1), 53–59.
26. Zenhausern, F.; O'Boyle, M.P.; Wickramasinghe, H.K. Apertureless near-field optical microscope. Appl. Phys. Lett. **1994**, *65* (13), 1623–1625.
27. Knoll, B.; Keilmann, F. Infrared conductivity mapping for nanoelectronics. Appl. Phys. Lett. **2000**, *77* (24), 3980–3982.
28. Yin, X.; Fang, N.; Zang, X.; Martini, I.B.; Schwartz, B.J. Near-field two-photon nanolithography using an apertureless optical probe. Appl. Phys. Lett. **2002**, *81* (19), 3663–3665.
29. Hillenbrand, R.; Keilmann, F. Complex optical constants on a subwavelength scale. Phys. Rev. Lett. **2000**, *85* (14), 3029–3032.
30. Drews, D.; Ehrfeld, W.; Lacher, M.; Mayr, K.; Noell, W.; Schmitt, S.; Abraham, M. Nanostruc-

tured probes for scanning near-field optical microscopy. Nanotechnology **1999**, *10*, 61–64.
31. Belier, B.; Santoso, A.; Bonnafe, J.; Nicu, L.; Temple-Boyer, P.; Bergaud, C. Micro-optomechanical sensor for optical connection in the near field. Appl. Phys. Lett. **2000**, *77* (12), 1768–1770.
32. Heisig, S.; Rudow, O.; Oesterschulze, E. Optical active gallium arsenide cantilever probes for combined scanning near-field optical microscopy and scanning force microscopy. J. Vac. Sci. Technol., B **2000**, *18* (3), 1134–1137.
33. Danzebrink, H.U.; Ohlsson, O.; Wilkening, G. Fabrication and characterization of optoelectronic near-field probes based on an SFM cantilever design. Ultramicroscopy **1995**, *61*, 131–138.
34. Tanaka, Y.; Fukuzawa, K.; Kuwano, H. Microfabrication of microtip on photocantilever for near-field scanning microscopy and investigation of effect of microtip shape on spatial resolution. J. Appl. Phys. **1998**, *83* (7), 3547–3551.
35. Davis, R.C.; Williams, C.C. Nanometer-scale absorption spectroscopy by near-field photodetection optical microscopy. Appl. Phys. Lett. **1996**, *69* (9), 1179–1181.
36. Rosner, B.T.; Bork, T.; Agrawal, V.; van der Weide, D.W. Microfabricated silicon coaxial field sensors for near-field scanning optical and microwave microscopy. Sens. Actuators, A **2002**, *102*, 185–194.
37. Dragnea, B.; Preusser, J.; Szarko, J.M.; Leone, S.R.; Hinsberg, W.D. Pattern characterization of deep-ultraviolet photoresists by near-field infrared microscopy. J. Vac. Sci. Technol., B **2001**, *19* (1), 142–152.
38. Keilmann, F.; Merz, R. Far-Infrared Near-Field Spectroscopy of Two-Dimensional Electron Systems. In *Near Field Optics*; Kluwer Academic Publishers: Dordrecht, 1993; 317–324.
39. Keilmann, F. FIR microscopy. Infrared Phys. Technol. **1995**, *36* (1), 217–224.
40. Hunsche, S.; Koch, M.; Brener, I.; Nuss, M.C. THz near-field imaging. Opt. Commun. **1998**, *150*, 22–26.
41. Lahrech, A.; Bachelot, R.; Gleyzes, P.; Boccara, A.C. Infrared-reflection-mode near-field microscopy using an apertureless probe with a resolution of 600. Opt. Lett. **1996**, *21* (17), 1315–1317.
42. Hillenbrand, R.; Knoll, B.; Keilmann, F. Pure optical contrast in scattering-type scanning near-field microscopy. J. Microsc. **2001**, *202* (1), 77–83.
43. Labardi, M.; Patane, S.; Allegrini, M. Artifact-free near-field optical imaging by apertureless microscopy. Appl. Phys. Lett. **2000**, *77* (5), 621–623.
44. Knoll, B.; Keilmann, F. Near-field probing of vibrational absorption for chemical microscopy. Nature **1999**, *399*, 134–137.
45. Hillenbrand, R.; Taubner, T.; Keilmann, F. Phonon-enhanced light–matter interaction at the nanometre scale. Nature **2002**, *418*, 159–162.
46. Grober, R.D. High Efficiency Near-Field Electromagnetic Probe Having a Bowtie Antenna Structure, US Patent #5,696,372, 1997.
47. Rosner, B.; Peck, J.; van der Weide, D.W. Near field antennas integrated with scanning probes for THz to visible microscopy: Scale modeling and limitations on performance. IEEE Trans. Antennas Propag. **2002**, *50*, 670–675.
48. van der Weide, D.W.; Keilmann, F. Coherent Periodically Pulsed Radiation Spectrometer, US Patent #5,748,309, 1998.
49. Bozec, L.; Hammiche, A.; Pollock, H.M.; Conroy, M.; Chalmers, J.M.; Everall, N.J.; Turin, L. Localized photothermal infrared spectroscopy using a proximal probe. J. Appl. Phys. **2001**, *90*, 5159–5165.
50. Hammiche, A.; Pollock, H.M.; Reading, M.; Claybourn, M.; Turner, P.; Jewkes, K. Photothermal FTIR spectroscopy: A step towards FT-IR microscopy at a resolution better than the diffraction limit. Appl. Spectrosc. **1999**, *53* (7), 810–815.
51. Kölzer, J.; Oesterschulze, E.; Deboy, G. Thermal imaging and measurement techniques for electronic materials and devices. Microelectron. Eng. **1996**, *31*, 251–270.
52. Dragnea, B.; Leone, S.R. Advances in submicron infrared vibrational band chemical imaging. Int. Rev. Phys. Chem. **2001**, *20* (1), 59–92.
53. Pollock, H.M.; Smith, D.A. The Use of Near-Field Probes for Vibrational Spectroscopy and Photothermal Imaging. In *Handbook of Vibrational Spectroscopy*; John Wiley & Sons: Indianapolis, 2002.
54. Keilmann, F.; Weide, D.W.v.d; Eickelkamp, T.; Merz, R.; Stöckle, D. Extreme sub-wavelength resolution with a scanning radio-frequency transmission microscope. Opt. Commun. **1996**, *129*, 15–18.
55. van der Weide, D.W. Localized picosecond resolution with a near-field microwave/scanning-force microscope. Appl. Phys. Lett. **1997**, *70* (6), 677–679.
56. Rosner, B.T.; van der Weide, D.W. High-frequency near-field microscopy. Rev. Sci. Instrum. **2002**, *73* (7), 2505–2525.
57. Heisig, S.; Oesterschulze, E. Gallium arsenide probes for scanning near-field probe microscopy. Appl. Phys., A **1998**, *66*, S385–S390.
58. Kramer, A.; Keilmann, F.; Knoll, B.; Guckenberger,

R. The coaxial tip as a nano-antenna for scanning near-field microwave transmission microscopy. Micron **1996**, *27* (6), 413–417.
59. McCarty, G.S.; Weiss, P.S. Scanning probe studies of single nanostructures. Chem. Rev. **1999**, *99* (7), 1983–1990.
60. Michel, B.; Mizutani, W.; Schierle, R.; Jarosch, A.; Knop, W.; Benedickter, H.; Bächthold, W.; Rohrer, H. Scanning surface harmonic microscopy: Scanning probe microscopy based on microwave field-induced harmonic generation. Rev. Sci. Instrum. **1992**, *63*, 4080–4085.
61. Stranick, S.J.; Bumm, L.A.; Kamna, M.M.; Weiss, P.S. Linear and Nonlinear Spectroscopy with the Tunable AC Scanning Tunneling Microscope. In *Photons and Local Probes*; Kluwer Academic Publishers: Dordrecht, 1995; 221–233.
62. Keilmann, F.; Knoll, B.; Kramer, A. Long-wave-infrared near-field microscopy. Phys. Status Solidi, B **1999**, *215*, 849–854.
63. Weiss, S.; Ogletree, D.F.; Botkin, D.; Salmeron, M.; Chemla, D.S. Ultrafast scanning probe microscopy. Appl. Phys. Lett. **1993**, *63* (18), 2567–2569.
64. Nunes, G.; Freeman, M.R. Picosecond resolution in scanning tunneling microscopy. Science **1993**, *262*, 1029–1032.
65. Botkin, D.; Glass, J.; Chemla, D.S.; Ogletree, D.F.; Salmeron, M.; Weiss, S. Advances in ultrafast scanning tunneling microscopy. Appl. Phys. Lett. **1996**, *69* (9), 1321–1323.
66. Kim, J.; Williamson, S.; Nees, J.; Wakana, S.; Whitaker, J. Photoconductive sampling probe with 2.3-ps temporal resolution and 4V sensitivity. Appl. Phys. Lett. **1993**, *62* (18), 2268–2270.
67. Takeuchi, K.; Mizuhara, A. Image observation of picosecond electrical pulse by scanning force optoelectronic microscope. IEEE MTT-S Int. Microw. Symp. Dig. **1997**, *3*, 1643–1646.
68. Oesterschulze, E.; Heisig, S.; Steffens, W.M. Cantilever probes for spatio-temporal imaging of voltage pulses with an ultrafast scanning probe microscope. J. Vac. Sci. Technol., B **2001**, *19* (1), 107–110.
69. Golosovsky, M.; Davidov, D. Novel millimeter-wave near-field resistivity microscope. Appl. Phys. Lett. **1996**, *68* (11), 1579–1581.
70. Abu-Teir, M.; Golosovsky, M.; Davidov, D.; Frenkel, A.; Goldberger, H. Near-field scanning microwave probe based on a dielectric resonator. Rev. Sci. Instrum. **2001**, *72* (4), 2073–2079.
71. Steinhauer, D.E.; Vlahacos, C.P.; Dutta, S.; Wellstood, F.C.; Anlage, S.M. Surface resistance imaging with a scanning near-field microwave microscope. Appl. Phys. Lett. **1997**, *71* (12), 1736–1738.
72. Tabib-Azar, M.; Su, D.P.; Pohar, A.; LeClair, S.R.; Ponchak, G. 0.4 μm spatial resolution with 1 GHz (λ=30 cm) evanescent microwave probe. Rev. Sci. Instrum. **1999**, *70* (3), 1725–1729.
73. Gao, C.; Wei, T.; Duewer, F.; Lu, Y.; Xiang, X.D. High spatial resolution quantitative microwave impedance microscopy by a scanning tip microwave near-field microscope. Appl. Phys. Lett. **1997**, *71* (13), 1872–1874.
74. Gao, C.; Duewer, F.; Lu, Y.; Xiang, X.D. Quantitative nonlinear dielectric microscopy of periodically polarized ferroelectric domains. Appl. Phys. Lett. **1998**, *73* (8), 1146–1148.
75. Duewer, F.; Gao, C.; Takeuchi, I.; Xiang, X.D. Tip–sample distance feedback control in a scanning evanescent microwave microscope. Appl. Phys. Lett. **1999**, *74* (18), 2696–2698.
76. Takeuchi, I.; Wei, T.; Duewer, F.; Yoo, Y.K.; Xiang, X.D.; Talyansky, V.; Pai, S.P.; Chen, G.J.; Venkatesan, T. Low temperature scanning-tip microwave near-field microscopy of $YBa_2Cu_3O_{7-x}$ films. Appl. Phys. Lett. **1997**, *71* (14), 2026–2028.
77. Golosovsky, M.; Galkin, A.; Davidov, D. High-spatial resolution resistivity mapping of large-area YBCO films by a near-field millimeter-wave microscope. IEEE Trans. Microw. Theor. Tech. **1996**, *44* (7), 1390–1392.
78. Tabib-Azar, M.; Ciocan, R.; Ponchak, G.; LeClair, S.R. Transient thermography using evanescent microwave microscope. Rev. Sci. Instrum. **1999**, *70* (8), 3387–3390.

Near-Field Raman Spectroscopy

Eric Ayars
California State University—Chico, Chico, California, U.S.A.

INTRODUCTION

Near-field Raman Spectroscopy (NFRS) combines the chemical-information-gathering power of Raman spectroscopy with the spatial resolution of the atomic force microscope (AFM). In addition, it allows the researcher to correlate topographic and optical data in ways impossible via other techniques. This article presents an overview of the methods used in NFRS, a brief history of the field, and a tour of some recent achievements in this rapidly evolving field. As with any other field at this stage of development, there are still unanswered questions in NFRS; some of these will be addressed as well.

NEAR-FIELD SCANNING OPTICAL MICROSCOPY

Microscopy is one of the most powerful tools available in modern science. Much of our basic understanding of the world comes from our ability to see the details of how things operate. Until recently, however, the available resolution for optical microscopes has been limited by diffraction effects arising from the wave nature of light. Near-field scanning optical microscopy (NSOM) provides a means of surpassing those limits, thus bringing to the researcher a variety of research tools available in visible-light optics at resolutions previously unattainable at those wavelengths. Because near-field Raman microscopy (NSRM) obtains both its resolution and its enhancement effects from the characteristics of light in the near field, it is necessary to begin any discussion of NSRM with an overview of NSOM.

Resolution Limits to Conventional Microscopy

The fundamental limit to resolution of conventional optical microscopes arises from the nature of light. Because light is a wave, it diffracts on interaction with the sample being studied. The diffraction angle depends on the light frequency and the spatial frequency of the features on the sample. In order to resolve an object, it is necessary to collect not only the central zeroth-order diffracted light, but some of the higher orders as well.[1] These higher-order diffraction maxima contain the spatial information.

As an example, consider a one-dimensional sample with some contrast mechanism having two distinct spatial frequencies (Fig. 1). Here, the contrast is represented as height; it could be contrast in color, refractive index, reflectivity, optical density, fluorescence, or anything else observable with light. The lens shown in Fig. 1 has sufficient numeric aperture (NA) to capture the light diffracted from the larger features, but light diffracted from the smaller features is not collected. Thus, the image formed does not include the components with higher spatial frequency. A lens with a higher NA could collect the light from these smaller features, but there is a practical limit to how far this sort of improvement can go. The highest NA obtainable with oil-immersion lenses is just over 1.4, resulting in a minimum resolution of roughly 250 nm for visible light.

How to Avoid Resolution Limits

NSOM allows higher resolution than would be allowed by the Abbe limit. It does this by utilizing the continuity of the fields near an aperture or sharpened conductor. The illumination pattern cannot change instantaneously from subwavelength source spot to broad far-field spot. In the far field, defined as distances greater than a few tip radii from the source, light is diffracted at extremely high angles and illuminates a broad region. At the tip, the light is constrained to a region near the tip.[2] Close to the tip, in the near field, the light is still limited to a relatively small region, on the order of the tip radius. By holding the sample within this near-field region, the illuminated area on the sample is kept small—smaller than would be allowed by the Abbe limit. Light interacting with the sample can be collected in the far field with conventional optics. Although the collection optics are incapable of resolving that illuminated region, the light collected is known to have originated from a well-defined small spot; by scanning the aperture over the sample one can form a rasterized image with optical resolution approximately equal to that spot size.

In the region near the tip, often called the near-field enhancement region, the electric field decreases exponentially with distance. This is of particular importance in

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013824

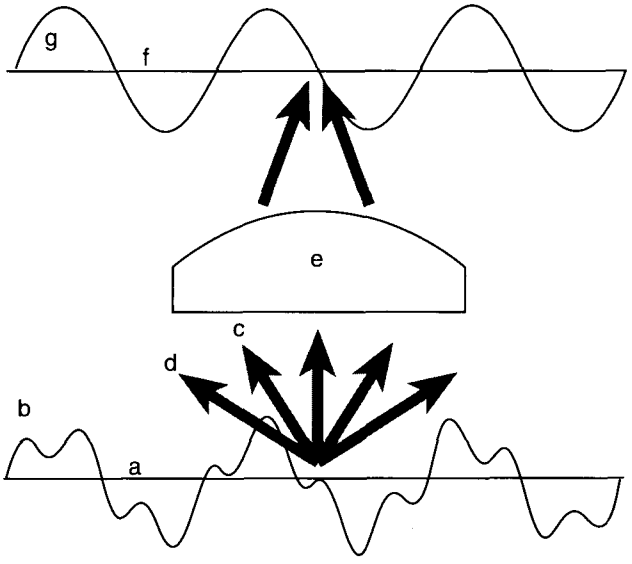


Fig. 1 The origin of the resolution limit in conventional microscopy. Object (a) has a contrast profile (b) with two distinct spatial frequencies. Light from the sample is diffracted by the larger features at angle (c), and by the smaller features at angle (d). The lens (e) collects only the light diffracted by the larger features, so the resulting image (f) contains information about only the larger features (g). Increasing the magnification would create a larger image of (g), with no increase in information content.

Raman applications, where the enhanced electric field in the near field can result in significantly increased Raman signal and may be responsible for other effects as well.

The earliest NSOM systems used a sharpened optical fiber as a tip. This was prepared by either heating and pulling or boundary-layer chemical etching[3] and then was coated with metal to constrain the electric field to the interior of the tip. An aperture at the end of this tip is formed by focused ion-beam machining or, more commonly, by holding the tip at an angle during the metal-coating process so that the end was shadowed and not coated. This subwavelength aperture is held close to the sample and acts as a light source or detector, or both, depending on the experimental configuration. More recent systems have used a sharpened metallic tip or metal-coated AFM tip as a scattering source. This "aperture-less" tip is held near the sample as before, and then illuminated externally.

Behavior of Light Near a Tip

The problem of diffraction through a subwavelength aperture was first solved by Bethe[4] and later improved by Bouwkamp.[5] Their solution assumes a circular aperture

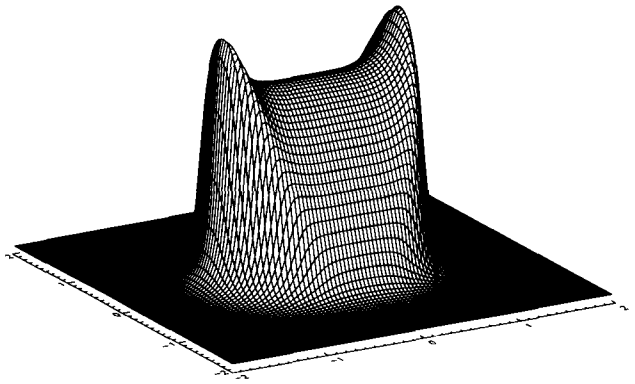


Fig. 2 Near-field enhancement. Graph height represents the intensity of the light emerging from a circular subwavelength aperture with the incident (x) polarization, with an aperture size a equal to one-fifth the wavelength, at a distance of one-tenth a.

in an infinitely thin, flat, perfect conductor. Other work has been done, e.g., with numeric calculations of the field components near a two-dimensional tip (a slit) with a more realistic tip profile made of real materials[6] and apertureless NSOM.[7–9] All of these calculations show the same basic characteristics in the resulting fields: very strong distance dependence, evanescent fields near the tip, and dependence on polarization.

The square of the electric field for light emerging from a tip aperture with the original polarization (Bethe–Bouwkamp model) is shown in Fig. 2. Here the aperture radius is one-fifth the wavelength of the light, and the electric field magnitude is calculated at a distance of one-tenth the aperture radius. Some light emerging from the aperture is polarized at 90° to the incident light; this

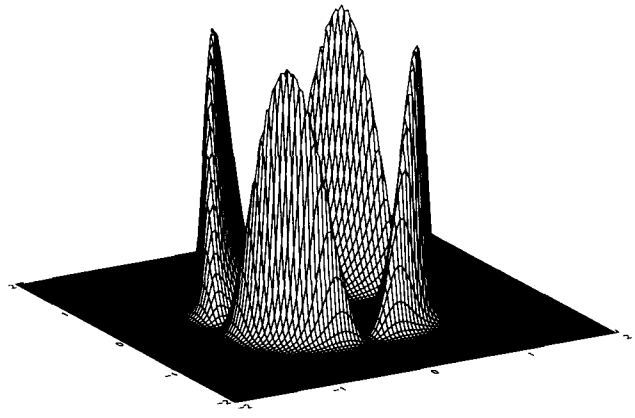


Fig. 3 Near-field enhancement, showing the component of the light intensity polarized at 90° (y) to the incident light. The aperture parameters are the same as in Fig. 2.

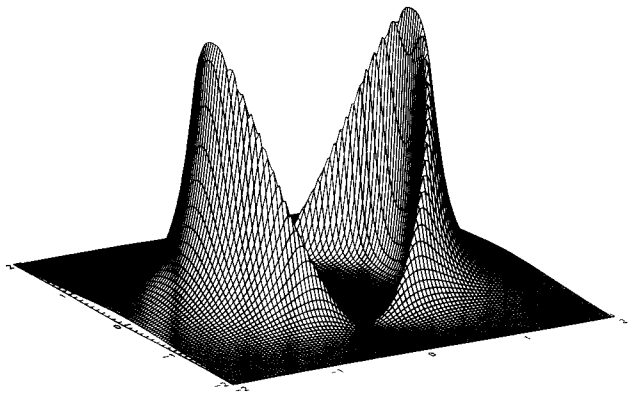


Fig. 4 Near-field enhancement, showing the component of the light intensity polarized in the direction of the incident light (z). The aperture parameters are the same as in Fig. 2.

profile (with the same tip parameters) is shown in Fig. 3. There is also a relatively strong component of the emerging light that is polarized in the direction of propagation, which is shown in Fig. 4. All of these fields, with the exception of the central component of the first, die out rapidly with increasing distance from the tip. Fig. 5 shows the field profile at increasing distance from the tip for the first case, corresponding to Fig. 2.

For apertureless tips, there is a similar enhancement region near the tip. Typically, the tip is modeled as a conical metal probe with a spherical apex. The enhancement region extends over a region approximately the size of the diameter of the tip apex, and only the component of the light polarized in the direction of the tip axis is strongly enhanced. This enhancement occurs for both metallic and dielectric probes, but for metallic probes there is additional enhancement because of surface plasmon–polaritons at the apex of the tip.[9] One significant advantage to apertureless tips is that the incident light intensity is not subject to the limitations[10] imposed by the threat of thermal breakdown of the reflective coating on a fiber optic tip. Ordinary lenses are used to focus the light on the tip, either from the side or through the sample from below with a high-NA lens.[11] The illuminated region is much larger than the tip; but the signal from the enhancement region at the tip is by far the dominant component of the collected signal.

Tip–Sample Distance Regulation

For either aperture or apertureless NSOM and NSRM, it is necessary to hold the tip at a close and well-regulated distance from the sample. The feedback mechanism responsible for maintaining this near-field distance is central to any microscope head design and depends on the type of tip used. Frequently, tips are formed by coating commercial cantilever AFM tips with metal. If this is the case, it is simplest to use standard AFM feedback methods, described elsewhere. If the tips used are formed from optical fibers or etched metal wires, or if it is desirable to control the tip–sample distance more precisely than is possible with relatively flexible cantilever AFM tips, then shear-force feedback must be used. In shear-force feedback, the tip is shaken, at or near its resonance, in a plane perpendicular to the plane of the sample. Some measurement of the tip oscillation amplitude is made, and as the tip approaches the sample this amplitude decreases. The feedback electronics keep the tip positioned via piezoelectric actuators so that the amplitude remains at some intermediate value between its free-oscillation value and zero. By recording the adjustment necessary to maintain a constant tip–sample distance, it is possible to map the topographic features of the sample. This topographic detail can then be correlated to the optical information for a better understanding of the phenomena being observed.

The earliest method of shear-force tip–sample distance control was optical feedback. In an optical feedback system, a laser beam is focused on the oscillating fiber near the tip. The diffraction pattern "shadow" of the tip falls on an edge-masked detector. The signal from this detector serves as a measurement of the tip oscillation amplitude, and thus provides a shear-force feedback signal.

In recent years, optical feedback has been largely superseded by tuning-fork feedback.[12] In the tuning-fork method for distance regulation, the optical fiber or wire,

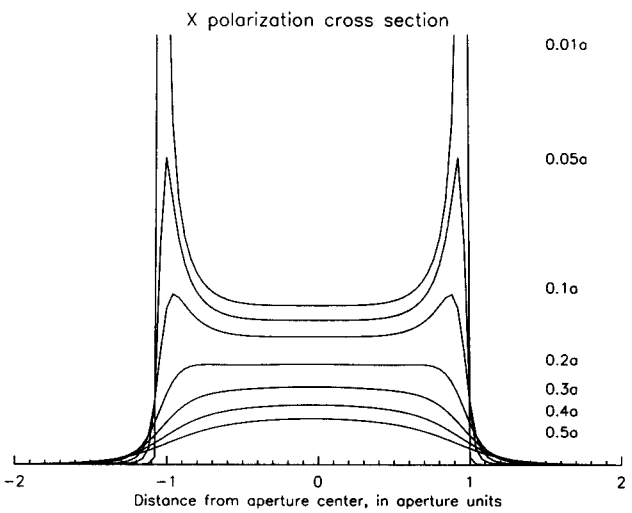


Fig. 5 Line cuts showing distance dependence of near-field enhancement. The line at $0.1a$ is a cut through the center of Fig. 2; the others show how that enhancement changes with distance.

with the subwavelength aperture or tip at the end, is glued to the side of a small quartz tuning fork. The fork/tip assembly is shaken at or near its resonance, and the piezoelectric tuning fork provides an electric signal proportional to the amplitude of oscillation of the tip. The tuning-fork method offers considerable advantages over the optical feedback method. It eliminates the difficulty of aligning and focusing the feedback laser with the fiber and edge mask. With optical feedback, any drift in feedback laser power or alignment of the feedback optics can be misinterpreted as a change in feedback level, causing problems in low-light experiments requiring long integration times. Tuning-fork feedback provides a feedback signal that is much more stable on long scans. Finally, tuning-fork feedback does not require a second laser beam, thus eliminating one more source of noise in spectroscopic or fluorescence applications.

RAMAN SPECTROSCOPY

Raman spectroscopy, discovered in 1928 by Chandrasekhara Venkata Raman, is the spectroscopy of vibrational transitions measured by the inelastic scattering of light. The energy of vibrational levels typically corresponds to wave numbers of 10^2 to 10^4 cm^{-1}, compared with the 10^4 to 10^6 cm^{-1} of valence electron transitions seen with visible and UV spectroscopy. The light wavelength required for absorption spectroscopy in this energy range is in the infrared to far-infrared range. This range of wavelengths is inconvenient to work with: appropriate optical materials for this entire range of wavelengths are scarce or nonexistent, requiring different gratings and detectors for different portions of the region, entirely reflective optics, and relatively large sample volumes. By measuring these transitions with inelastic scattering, it is possible to cover the entire vibrational spectrum with a single recording using easily managed visible light. In addition, the selection rules for IR spectroscopy and for Raman spectroscopy are markedly different. Some vibrations are strongly IR active but weak or nonexistent in Raman; others are strongly Raman active but weak in IR. The two tools should be regarded as complementary: For a complete understanding of the vibrational levels of the material, it is often necessary to study both IR and Raman spectra.

Raman spectroscopy has advantages over IR spectroscopy. The primary advantage is the ability to do the spectroscopy with visible light. The benefits of using conventional optics, readily available light sources, and visible-light alignment are not to be underestimated. Another advantage is the ability to work with small sample volumes. Micro-Raman systems typically can work with a sample region of about 1-μm diameter on a surface, whereas IR spectroscopy, as an absorptive phenomenon, requires a larger volume of material to have a significant effect. Water is a strong absorber of IR light, so it is difficult to do IR spectroscopy of compounds in aqueous solution, but water is a weak Raman scatterer, making it ideal for the study of biological materials in solution.

In the classical one-dimensional derivation of the Raman effect,[13] the polarization P of the material depends on the polarizability α of the material and on the incident electric field E, which is provided by the incident light at frequency ν_o:

$$P = \alpha E_o \cos(2\pi\nu_o t) \tag{1}$$

The electric field magnitude E_o is presumed to be constant over the dimensions of the oscillation. This is an excellent approximation for visible light in the far field. Charges in the material may oscillate at some frequency ν_m about their equilibrium positions:

$$q = q_o \cos(2\pi\nu_m t) \tag{2}$$

This oscillation may induce a change in the polarizability of the material, which can be approximated by a Taylor expansion of α:

$$\alpha \approx \alpha_o + \left(\frac{\partial\alpha}{\partial q}\right) q_o + \cdots \tag{3}$$

Combining these and simplifying gives, to first order in q:

$$P \approx \alpha_o E_o \cos(2\pi\nu_o t) + \frac{1}{2} q_o \left(\frac{\partial\alpha}{\partial q}\right)_o E_o[\cos\{2\pi(\nu_o - \nu_m)t\} + \cos\{2\pi(\nu_o + \nu_m)t\}] + \cdots \tag{4}$$

The first of these terms is the Rayleigh scattering, which is unshifted in frequency. The second is the Raman scattering, at the characteristic Stokes (−) and anti-Stokes (+) frequencies. The existence of Raman-scattered light depends on the Raman activity, $\partial\alpha/\partial q$: if there is no change in the polarizability for a given vibrational transition, $\partial\alpha/\partial q = 0$ and the transition is not Raman active.

In three dimensions, the equations become tensorial, but other than that the derivation is essentially unchanged. Change in the polarizability tensor α, which is symmetric for conventional Raman scattering, still determines the activity, but now the vibration is Raman active if any one of the polarizability tensor components is changed during the vibration. Additionally, the three-dimensional form of the derivation brings to light the polarization dependence of the Raman activity. It is possible, in crystalline materials, e.g., to have normal modes of vibration that

are excited only by certain polarization states of the incident light. This is a great advantage in solid-state studies, because it allows investigation of anisotropic bonds for a more complete understanding of the material.

The primary experimental difficulty with Raman spectroscopy is signal-to-noise ratio. The quantum efficiency of Raman spectroscopy is typically on the order of 10^{-6}: for every million photons that strike the sample, one might be Raman shifted. This low signal level is exacerbated by the close proximity, spectrally, of the Rayleigh-scattered light. The spectrometer must be designed to minimize internal scattering of this unshifted light, as even slight scattering will overwhelm the small Raman signal. A double spectrometer has long been the accepted minimum for Raman work, and triple spectrometers are not unusual. These cut the amount of light that reaches the detector, but they cut the amount of scattered light that reaches the detector even more. In recent years, the development of holographic notch filters has allowed researchers to block the unshifted light before it reaches the spectrometer.[14–16] This allows use of single spectrometers and lessens the signal losses associated with double and triple spectrometers. Use of a holographic notch filter and single spectrometer also allows one to collect the final signal with a charge-coupled device (CCD). This permits collection of an entire spectrum at once, rather than one point at a time, which is a great advantage when—as in NSRM—long integration times are necessary.

NEAR-FIELD RAMAN SPECTROSCOPY

The primary difficulty with Raman spectroscopy is also the primary weakness of near-field scanning optical microscopy: low signal levels. The amount of light that can be sent through a subwavelength probe aperture is severely limited by the shape and material of the probe.[17] The combination of this low excitation power and the low cross section for Raman scattering necessitates long integration times. In order to image the sample, this long integration time must be applied at each point in the scan.

The first NSOM image successfully taken on a Raman line was achieved in 1995.[18] The image, a 4-μm^2 square area on a sample of Rb-doped potassium titanyl phosphate

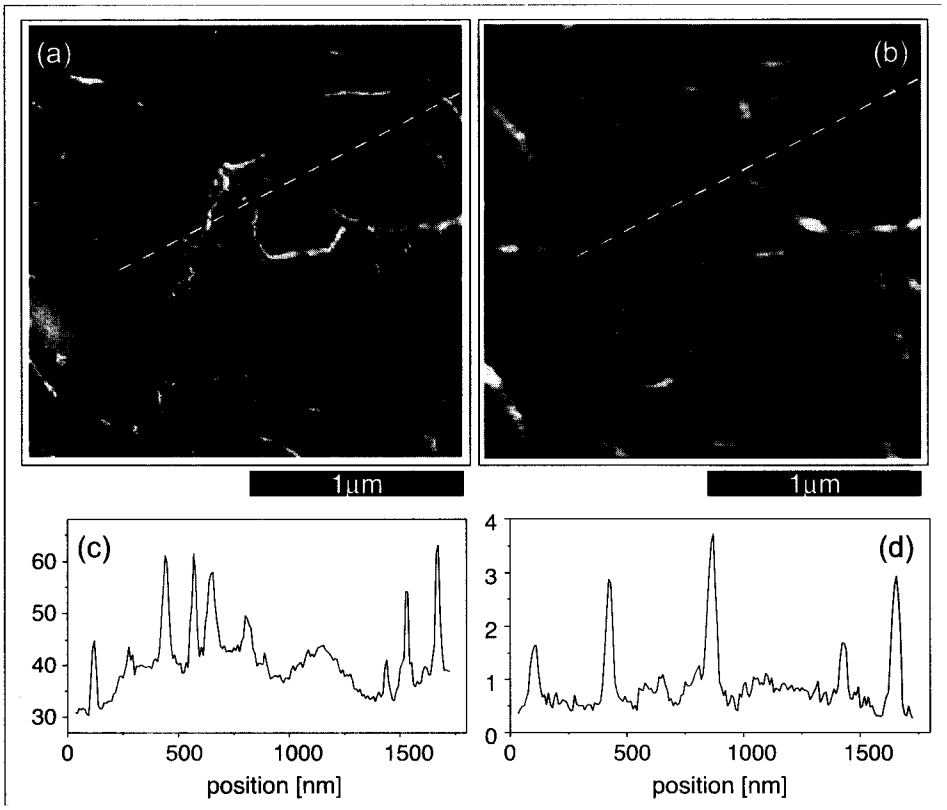


Fig. 6 Raman image of carbon nanotubes (a) together with simultaneously acquired topography (b). Line scans (c) and (d) are taken along the dotted lines in (a) and (b), and have vertical scales of kcts/sec and nm, respectively. (Courtesy of Achim Hartschuh, Physikalische Chemie, Universität Siegen.)

(KTP), used the difference in Raman spectra between the pure and the doped regions as a contrast mechanism. It used a tapered fiber probe, had a resolution of approximately 250 nm, and took over 10 hr to complete.

Apertureless probes offer much higher illumination power, and may offer greater near-field enhancement as well.[a] The resolution of an NSOM image is dependent on the size of the probe, so the smaller tip size for an apertureless probe also offers a resolution advantage. The highest-resolution NSRM images published to date, with spatial resolution of less than 25 nm, have been produced using silver apertureless probes[19] (Fig. 6). The data represented by Fig. 6 took approximately 40 min to collect.

The significant decrease in the time it takes to obtain an NSRM image with apertureless probes has resulted in their almost universal adoption. It has also sparked a surge in research applications. Recent uses of NSRM include measurement of residual stress in a damaged silicon wafer,[20] imaging of organic molecules,[21] and measurement of variation in Raman spectra along single-walled nanotubes[19] as well as many others.

Differences Between NSRM and Micro-Raman

It is tempting to view NSOM-Raman spectroscopy as merely the next resolution increase in a well-understood chain of improvements. Micro-Raman spectroscopy increased the spatial resolution of Raman studies far beyond that of ordinary Raman work,[22] and on a superficial level, NSOM-Raman is seen as the next step in resolution enhancement. This is not entirely the case: there are aspects of NSRM that are fundamentally different from micro-Raman.

The differences between NSRM and other Raman spectroscopies originate with the nature of light in the near field. As discussed above, the electric field components near an NSOM tip are not uniform in either space or polarization. For an apertured tip, with x-polarized incident light, the transmitted light has x, y, and z components. With apertureless tips, there is a similarly strong enhancement of the polarization component along the tip axis. The total intensity of the radiation varies strongly with distance from the tip. This variation is due to the presence, in the near field, of evanescent (nonpropagating) modes, which exist for each of the three polarization states.

[a]The magnitude of the enhancement is inversely related to the radius of curvature of the probe. Apertureless probes are typically much sharper than tapered fiber probes, although the geometry of the two is different enough that a direct comparison is difficult.

The first of the effects in near-field Raman spectroscopy is due to the overall enhancement of the light intensity near the probe aperture. This enhancement does not extend to the far field, and thus does not contribute to measurements of the total radiation emitted from the probe. Because this total intensity near the probe tip includes the evanescent modes, which do not contribute to any far-field intensity measurements, there is actually more light available for Raman excitation in the near field than one would expect. This effect has proven useful in measuring the Raman spectra of extremely thin films: in the far field, there is no observed Raman signal from the film, but when a tip is in the near field there is a strong signal from the film.[23] The same effect may be useful for measurements of surface states on solids.

The z polarization component of the light near the probe also leads to unusual near-field Raman effects.[24] In conventional Raman spectroscopy, vibrational modes with x or y components are excited by the incident light, which is necessarily polarized in the xy plane. Modes with purely z components require an electric field in the z direction for excitation. In liquid- or gas-phase Raman work, this is not a limitation, likewise in amorphous materials or polycrystalline materials with domain size less than the minimum spot size. In single-crystal work, the z modes are usually measured by reorienting the crystal so that the incident light has a polarization component along the desired axis. In near-field Raman, there is a z component of the field that provides excitation of these modes without reorientation of the crystal. This effect can be effectively "turned off" by moving the tip out of the near field. Comparison of near-field and far-field spectra allows the researcher to determine the polarization dependence of the material.

The third of the effects possible in NSRM is due to the strong gradient in the near-field electric fields.[25] In the standard derivation of the Raman effect (Eqs. 1–4) it is assumed that the electric field magnitude remains constant over the entire extent of the oscillation. For conventional Raman techniques, this is true. Even for confocal Raman spectroscopy, this is an excellent approximation; the intensity of the traveling wave is not going to change measurably over molecular distances. With NSRM, however, the approximation no longer holds. The traveling-wave component of the transmitted radiation is only part of the oscillating field present in NSOM: the evanescent modes also contribute to the total electric field and to the Raman excitation.

The effect of the field gradient can be explored by taking a Taylor expansion of the electric field.

$$E \approx E_0 + \left(\frac{\partial E}{\partial q}\right)_0 q + \cdots \tag{5}$$

Inserting that expansion in place of E_o in the derivation given in Eqs. 1–4 gives:

$$P \approx \alpha_o E_o \cos(2\pi\nu_o t) + \frac{1}{2} q_o \left(\frac{\partial\alpha}{\partial q}\right)_o E_o[\cos\{2\pi(\nu_o - \nu_m)t\} + \cos\{2\pi(\nu_o + \nu_m)t\}] + \frac{1}{2} q_o \left(\frac{\partial E}{\partial q}\right)_o \alpha_o[\cos\{2\pi(\nu_o - \nu_m)t\} + \cos\{2\pi(\nu_o + \nu_m)t\}] + \cdots \quad (6)$$

The first two terms of Eq. 6 are the same as before. The third describes a scattering with the same Stokes and anti-Stokes frequency shifts as the Raman lines, but no dependence on the Raman activity $\partial\alpha/\partial q$. Instead, this scattering term is dependent on the field gradient and the polarizability. Selection rules for this term depend on the characteristics of α, and so would have activity similar to IR absorption spectroscopy. The net result would be the observation of transitions that are normally seen in IR absorption spectroscopy but are not Raman active.

The effect of a strong electric field gradient on Raman spectra has been discussed previously as a mechanism for some of the observed spectral lines in surface-enhanced Raman spectroscopy (SERS).[26] Because there is no good, controllable method of regulating distance or field gradient in SERS, NFRS offers the first opportunity to test the distance dependence of this effect. The effect was also predicted as an effect in the Raman spectra of microparticles suspended in laser traps.[27] This prediction was tested at the time, but for the gradients available in that experimental configuration the strength of the effect was calculated to be four orders of magnitude lower than the Raman effect and no lines were observed. There has been at least one possible sighting of this field-gradient effect in NFRS[28] although some other researchers who have looked for it have not seen it.[19]

CONCLUSION

Near-field Raman spectroscopy allows researchers to make measurements of chemical bonds and vibrational transitions with spatial resolution unavailable via conventional optical methods. These high-resolution spectroscopic data come with corresponding topographic information, allowing better understanding of how nanoscale features affect the nanoscale spectroscopic data. The enhancement region near the tip falls off very rapidly, so NSRM can be used for Raman measurements of thin films and surface effects that would otherwise be either too small to observe or washed out by the bulk signal.

The primary difficulties in NSRM are low signal level and correspondingly long integration times. The development and use of apertureless tips has greatly increased the signal strength available, decreased the time required for image collection, and increased the resolution as well. These improvements have brought NSRM to the point where it is possible to obtain detailed high-resolution images, with separate topographic and Raman data, in well under an hour.

NSRM is not simply the next resolution step beyond micro-Raman. The presence of the metallic probe near the sample causes changes in intensity and polarization of the incident light, resulting in effects similar to surface-enhanced Raman spectroscopy. It is also possible, although not yet entirely certain, that the strong field gradients created by the tip may make visible some lines that are not normally Raman active.

REFERENCES

1. Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Arch. Mikrosk. Anat. **1873**, *9*, 413–468.
2. Betzig, E.; Trautman, J.K. Near-field optics: Microscopy, spectroscopy, and surface modification beyond the diffraction limit. Science **1992**, *257*, 189–195.
3. Hoffmann, P.; Dutoit, B.; Salathé, R.P. Comparison of mechanically drawn and protection layer chemically etched optical fiber tips. Ultramicroscopy **1995**, *61*, 165–170.
4. Bethe, H.A. Theory of diffraction by small holes. Phys. Rev. **1944**, *66* (7), 163–182.
5. Bouwkamp, C.J. Phillips Res. Rep. **1950**, *5*, 401.
6. Novotny, L.; Pohl, D.W.; Regli, P. Light propagation through nanometer-sized structures: The two-dimensional-aperture scanning near-field optical microscope. J. Opt. Soc. Am., A **1994**, *11* (6), 1768–1778.
7. Jersch, J.; Demming, F.; Hildenhagen, L.J.; Dickmann, K. Field enhancement of optical radiation in the nearfield of scanning probe microscope tips. Appl. Phys., A **1998**, *66* (1), 29–34.
8. Milner, R.G.; Richards, D. The role of tip plasmons in near-field Raman microscopy. J. Microsc. **2001**, *202*, 66–71.
9. Furukawa, H.; Kawata, S. Local field enhancement

with an apertureless near-field-microscope probe. Opt. Commun. **1998**, *148*, 221–224.
10. La Rosa, A.H.; Yakobson, B.I.; Hallen, H.D. Origins and effects of thermal processes on near-field optical probes. Appl. Phys. Lett. **1995**, *67* (18), 2597–2599.
11. Hayazawa, N.; Tarun, A.; Inouye, Y.; Kawata, S. Near-field enhanced Raman spectroscopy using side illumination optics. J. Appl. Phys. **2002**, *92* (12), 6983–6986.
12. Karrai, K.; Grober, R.D. Piezoelectric tip–sample distance control for near-field optical microscopes. Appl. Phys. Lett. **1995**, *66* (14), 1842–1844.
13. Ferraro, J.R.; Nakamoto, K. Basic Theory. In *Introductory Raman Spectroscopy*; Academic Press, Inc.: San Diego, 1994; 13–17.
14. Carrabba, M.; Spencer, K.; Rich, C.; Rauh, D. The utilization of a holographic Bragg diffraction filter for Rayleigh line rejection in Raman spectroscopy. Appl. Spectrosc. **1990**, *44* (9), 1558–1561.
15. Pelletier, M.J.; Reeder, R.C. Characterization of holographic band-reject filters designed for Raman spectroscopy. Appl. Spectrosc. **1991**, *45* (5), 765–770.
16. Sharma, S.K.; Schoen, C.L.; Helsley, C.E.; Owen, H. Performance of a holographic supernotch filter. Appl. Spectrosc. **1993**, *47* (3), 305–308.
17. Paesler, M.A.; Moyer, P.J. The Tapered Optical Fiber and Other Sensing Elements. In *Near-Field Optics: Theory, Instrumentation, and Applications*; John Wiley & Sons, Inc.: New York, 1996; 46–53.
18. Jahncke, C.L.; Paesler, M.A.; Hallen, H.D. Raman imaging with near-field scanning optical microscopy. Appl. Phys. Lett. **1995**, *67* (17), 2483–2485.
19. Hartschuh, A.; Sánchez, E.J.; Xie, X.S.; Novotny, L. High-resolution near-field Raman microscopy of single-walled carbon nanotubes. Phys. Rev. Lett. **2003**, *90* (9), 95503.
20. Webster, S.; Batchelder, D.N.; Smith, D.A. Submicron resolution measurement of stress on silicon by near-field Raman spectroscopy. Appl. Phys. Lett. **1998**, *72* (12), 1478–1480.
21. Hayazawa, N.; Inouye, Y.; Sekkat, Z.; Kawata, S. Near-field Raman imaging of organic molecules by an apertureless metallic probe scanning optical microscope. J. Chem. Phys. **2002**, *117* (3), 1296–1301.
22. Delhaye, M.; Dhamelincourt, P. Raman microprobe and microscope with laser excitation. J. Raman Spectrosc. **1975**, *3*, 33–43.
23. Anderson, M.S. Locally enhanced Raman spectroscopy with an atomic force microscope. Appl. Phys. Lett. **2000**, *76* (21), 3130–3132.
24. Ayars, E.J.; Hallen, H.D. Surface enhancement in near-field Raman spectroscopy. Appl. Phys. Lett. **2000**, *76* (26), 3911–3914.
25. Ayars, E.J.; Hallen, H.D.; Jahncke, C.L. Electric field gradient effects in Raman spectroscopy. Phys. Rev. Lett. **2000**, *85* (19), 4180–4183.
26. Sass, J.K.; Neff, H.; Moskovits, M.; Holloway, S. Electric field gradient effects on the spectroscopy of adsorbed molecules. J. Phys. Chem. **1981**, *85*, 621–623.
27. Knoll, P.; Marchl, M.; Kiefer, W. Raman spectroscopy of microparticles in laser light traps. Indian J. Pure Appl. Phys. **1988**, *26*, 268–277.
28. Ayars, E.J.; Jahncke, C.L.; Paesler, M.A.; Hallen, H.D. Fundamental differences between micro- and nano-Raman spectroscopy. J. Microsc. **2001**, *202* (1), 142–147.

Near-Field Raman Spectroscopy: Enhancing Spatial Resolution Using Metallic Tips

Satoshi Kawata
Yasushi Inouye
Osaka University, Osaka, Japan

INTRODUCTION

Optical microscopy is a well-established technology of microanalysis for material science, biology, pathological diagnostics, industry, and so on. However, the smallest analyzing dimension of the sample resolvable with a conventional optical microscope is a half wavelength of light or several hundreds of nanometers in visible region because of the nature of the light wave. This is a so-called diffraction limit of light, which is not small enough to observe the structures of interest in the current advanced sciences and technologies, including protein molecules, quantum nanodevices, and self-assembled molecules. Near-field optical microscopy is a tool to overcome the barrier of wavelength by confining photons in nanoscale volume.

A combination of the state-of-the-art near-field technology with the conventional optical measurements enables us to explore the nanoworld with visible light.[1,2] Especially, near-field Raman spectroscopy makes it possible to assign molecules, to analyze chemical behavior of molecules, and to observe molecular dynamics at nanometric or molecular scale.[3,4] Lasers and detectors of visible region are available in Raman spectroscopy, and quenching phenomenon[5] and photobleaching could be avoided; hence near-field Raman spectroscopy is suitable for molecular sensing with a nanometric spatial resolution. Because the cross section of Raman scattering is much smaller than fluorescence and infrared absorption cross section, it is necessary to enhance Raman cross section for its accurate measurement. This can be performed by using a metallic tip that also works as one of a near-field probe for detection. In this article, we present how the metallic tip works as a near-field probe at first and then show the current progress in near-field Raman spectroscopy using a metallic tip.

LOCAL FIELD ENHANCEMENT AT A METALLIC TIP

A small metallic structure with size smaller than the probing wavelength scatters light field and enhances the electric field intensely because of the strong interaction between the light field and the metal structure when the light is incident on such a structure. The reason for such an enhancement of the light field is that collective oscillation of free electrons in the metal resonates with the light field. This phenomenon is known as localized surface plasmon polaritons (SPPs). The electric field, which is coupled with the localized SPPs at the metallic structure, comprises evanescent field. Because the evanescent field is localized around the structure, the superresolving capability is attained by detecting the field for observation, measurement, and so on. Such a scheme of near-field scanning optical microscopy (NSOM) with a metallic probe of which the tip enhances electric field locally and strongly was first proposed in 1994.[6] The NSOM is called as an apertureless NSOM,[7,8] and the resolution is determined by the radius of the tip.

An apex of a metallic probe, such as a tip of a scanning tunneling microscopy (STM), works as a scatterer for enhancing electric field locally.[9] Fig. 1a shows intensity distribution of the light field scattered at the tip that is obtained using a numerical analysis.[1] Fig. 1b shows a schematic model for calculation. In this model, it is considered that a silver metallic tip with radius 20 nm is placed in contact on a glass substrate (refractive index: 1.5). The silver tip is illuminated with plane wave traveling in the substrate. Wavelength of the incident field is 488 nm, and its polarization is TM mode (i.e., p-polarization). Incident angle is 45°. Evanescent field is generated over surface of the glass substrate because the incident angle meets the condition of total internal reflection. Finite-difference time-domain (FDTD) method was used in the calculation.[10,11] Localized and enhanced field spot is observed in the proximity of the tip apex of Fig. 1a. The size of the small spot is around 30 nm, while approximately corresponds to the radius of the tip. Maximum intensity of the small spot was calculated to be 80 times as large as intensity of the incident field. Considering the tip to be a glass probe having the same tip radius of 20 nm, enhancement factor of the glass tip was calculated to be 7 when illuminating with TM mode. No enhancement of electric field was estimated at the metallic tip when TE mode (i.e., s-polarization) is used for illumination light in the calculation. These analyses show that

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009271

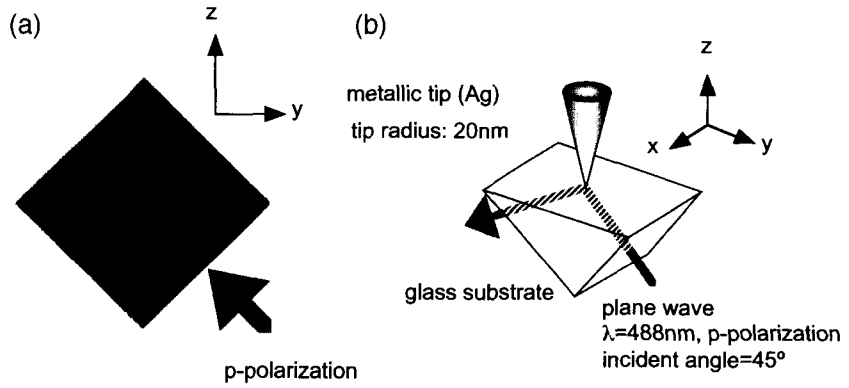


Fig. 1 Field enhancement effect induced by a metallic tip. (a) Intensity distribution of electric field locally enhanced at a silver probe tip calculated by the FDTD method. (b) Schematic of the calculation model for the analysis. (*View this art in color at www.dekker.com.*)

strong enhancement of the localized field requires the use of a metallic tip and TM mode illumination in an apertureless probe, and that the small light field spot strongly enhanced at the tip is caused by excitation of the localized SPPs.

NEAR-FIELD RAMAN SPECTROSCOPY AND MICROSCOPY

Fig. 2 shows illustrations of (a) near-field excitation of Raman scattering and (b) conventional Raman micro-spectroscopy (far-field detection). For near-field Raman spectroscopy, a metallic tip is placed in the vicinity of the specimen and local electric field is generated at the tip apex. Raman scattering of molecules, which are just below the enhanced electric field at the metallic tip, is selectively enhanced and detected with a sufficiently high spatial resolution.

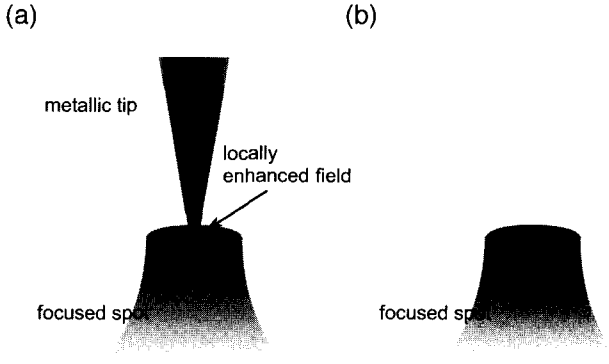


Fig. 2 Schematic illustration of illumination for (a) near-field Raman spectroscopy using a locally enhanced field, (b) conventional micro-Raman spectroscopy.

An optical setup of near-field Raman microscopy is shown in Fig. 3.[3,4,12,13] An expanded and collimated light field from a laser enters into the epi-illumination inverted microscope. A circular mask is inserted in the optical path of the illumination light and located at the conjugate plane of the pupil of the objective lens with numerical aperture (NA) equal to 1.4. Because the mask rejects part of the beam corresponding to focusing angles that are less than NA=1.0, the transmitted light forms a focused spot that produces an evanescent field on the sample surface.[14] As the metallic tip approaches to the focused spot, a locally enhanced electric field is observed to be generated at the tip apex. Fig. 4 shows CCD images of the scattering of the evanescent field at the metallic tip that were observed from the side of the tip. In Fig. 4a, the tip is in contact with the sample surface placed on the focused spot, which corresponds to Fig. 2a. Strong

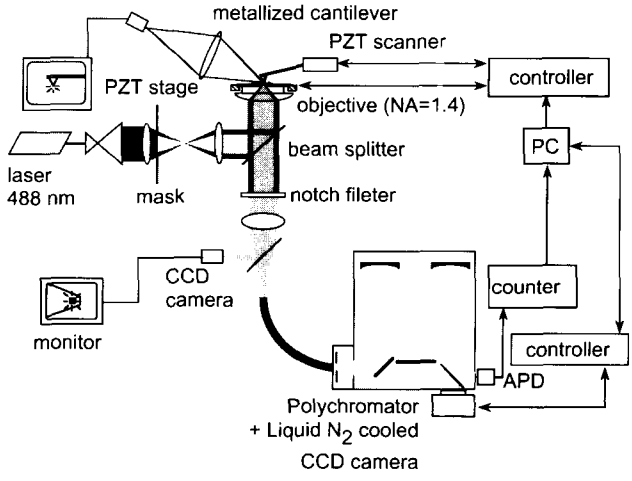


Fig. 3 Optical setup of near-field Raman spectroscopy using a metallic tip. (*View this art in color at www.dekker.com.*)

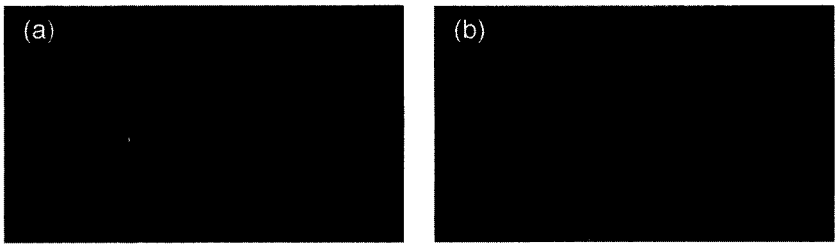


Fig. 4 (a) Photograph of a metallized cantilever tip shining in the focused evanescent field. (b) Photograph of a metallized tip which is retracted 500 nm from the sample.

Rayleigh scattering of the evanescent field at the tip apex is seen in the figure. In Fig. 4a and 4b, the tip is 0 and 500 nm away from the sample surface, respectively. No scattering of the evanescent field is observed because the tip is out of the skin depth of the evanescent field, while only very weak Rayleigh scattering light at the focused spot is observed because of the sample roughness. The electric field enhanced locally at the tip is scattered inelastically by Raman active molecules, which corresponds to near-field Raman scattering. Raman scattering is collected by the same objective lens and is directed to the spectrophotometer (focal length = 300 mm, 1200 lines/mm) that is equipped with a liquid-nitrogen-cooled CCD camera (1340 × 400 channels) for Raman spectra measurement and with an avalanche photodiode (APD) for Raman imaging. The APD is located after the exit slit of the spectrophotometer so that a specific Stokes-shifted line can be detected. Excitation light or Rayleigh scattering is sufficiently rejected by a notch filter. The metallic tip is a silicon cantilever that is coated with a 40-nm-thick silver film by a thermal evaporation process. The evaporation is performed at a relatively slow rate of 0.3 Å per second to avoid undesirable bending of the silicon lever. Fig. 5 shows a scanning electron micrograph of a metallized tip. Diameter of the silver-coated tip apex is around 40 nm. The distance between the sample and the cantilever is regulated in contact-mode AFM operation, and the sample is scanned with piezoelectric transducers (PZT) in the X–Y plane. Near-field Raman images are reconstructed by detecting near-field Raman signal at a specific Stokes-shifted line with the APD while scanning the PZT stage on the sample.

Raman spectra of Rhodamine 6G obtained with the near-field Raman microscope are shown in Fig. 6. Spectrum (a) is obtained with a silver-coated cantilever, and spectrum (b) is obtained without it. Intensity of spectrum (a) is higher than that of (b) because of the local field enhancement at the silver tip. Several Stokes-shifted Raman lines are observed in the spectrum. Acquisition time of the spectrum was 5 sec, and no accumulation was performed. The sample was made by casting the solution of Rhodamine 6G (6.0×10^{-4} wt) onto the 8-nm silver island-coated coverslip and depositing Rhodamine 6G because of evaporation of the solution. Raman scattering

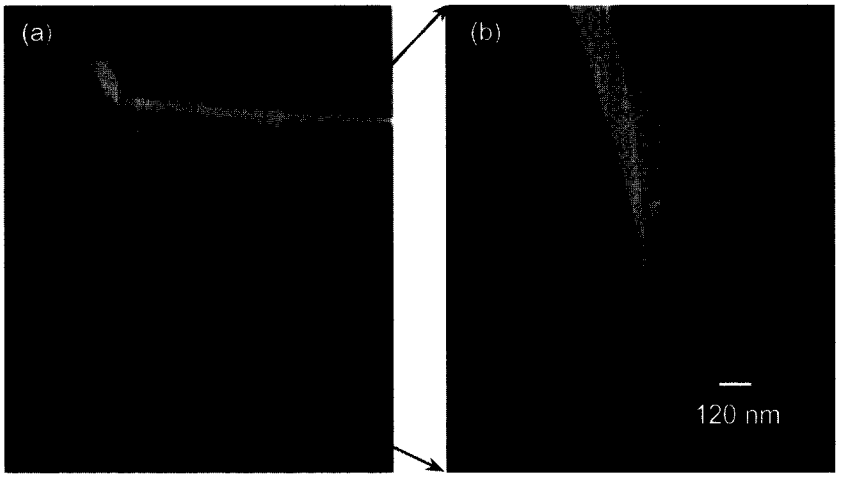


Fig. 5 Scanning electron microscopy images of the metallized cantilever tip. (a) Whole image of the tip. (b) Expanded image of the apex. (*View this art in color at www.dekker.com.*)

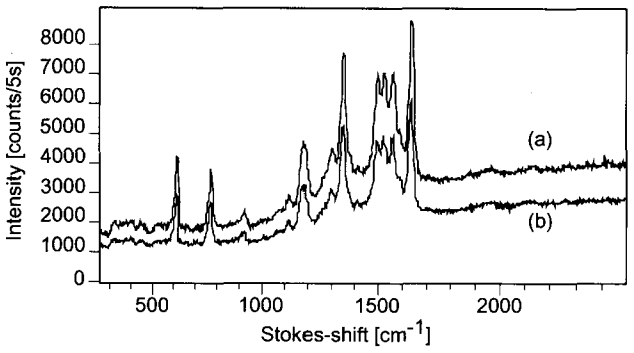


Fig. 6 Raman spectra of Rhodamine 6G. (a) Spectrum obtained with locally enhanced field. (b) Spectrum of conventional micro-Raman measurement.

is hence doubly enhanced because of surface-enhanced Raman scattering (SERS) both with the metallized tip and silver island film coated on substrate.

The Raman spectrum of Fig. 6a includes not only near-field component, but also far-field component. Near-field spectrum is obtained by subtracting the far-field spectrum of Fig. 6b obtained without the tip from the spectrum shown in Fig. 6a. Fig. 7 shows the result of subtraction or a near-field Raman spectrum of Rhodamine 6G. Each of Stokes-shifted peaks can be assigned to vibrational mode, such as peaks at 1359, 1503, 1570, and 1647 cm^{-1} which correspond to the aromatic C–C stretching vibrational modes, a peak at 1269 cm^{-1} which corresponds to the asymmetric vibrational mode of C–O–C bonding, a peak at 1120 cm^{-1} which corresponds to the in-plane bending mode of C–H bonding, and a peak at 766 cm^{-1} which corresponds to the out-of-plane bending mode of C–H bonding. Assignment of all Stokes-shifted peaks observed in Fig. 7 is shown in Table 1. When a silicon cantilever was used for near-field Raman detection, no enhancement is observed.[4]

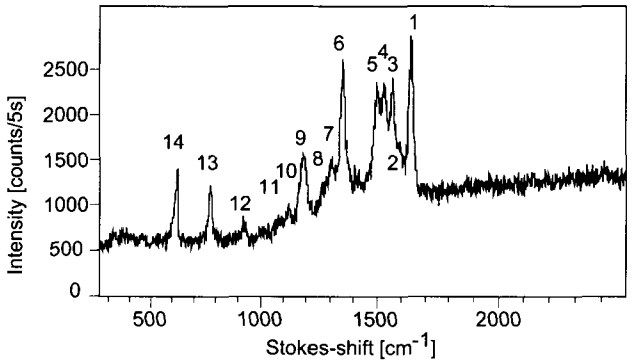


Fig. 7 Near-field Raman spectrum of Rhodamine 6G obtained by subtracting the spectrum of Fig. 6b from that of Fig. 6a.

Table 1 Assignment of Rhodamine 6G molecules

	Stokes shift (cm^{-1})	Assignment
1	1647	arom C–C str
2	1596	
3	1570	arom C–C str
4	1532	
5	1503	arom C–C str
6	1359	arom C–C str
7	1308	
8	1269	C–O–C
9	1185	
10	1120	C–H ip bend
11	1084	
12	919	
13	766	C–H op bend
14	608	C–C–C ip bend

arom: aromatic; bend: bending; str: stretching; ip: in-plane; op: out-of-plane.

Fig. 8 shows the Raman spectra of crystal violet molecules. Spectra (a) and (b) have been obtained with and without the metallic tip, respectively. The peaks at 1383, 1537, 1586, and 1617 cm^{-1} are assigned to the C–C

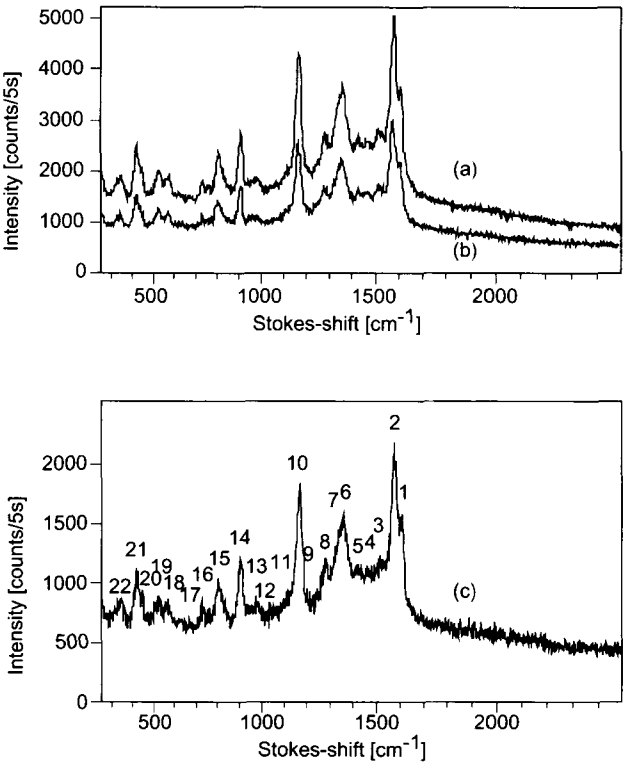


Fig. 8 Raman spectra of crystal violet. (a) Spectrum obtained with locally enhanced field. (b) Spectrum of conventional micro-Raman measurement. (c) Near-field Raman spectrum of crystal violet obtained by subtracting the spectrum of (b) from that of (a).

Table 2 Assignment of crystal violet molecules

	Stokes shift (cm^{-1})	Assignment
1	1617	arom C–C str
2	1586	arom C–C str
3	1537	arom C–C str
4	1475	arom C–C str
5	1440	arom C–C str
6	1383	arom C–C str
7	1363	*N*-phenyl
8	1290	
9	1211	
10	1173	C–H ip bend
11	1115	
12	980	
13	937	
14	908	C–H op bend
15	798	C–H op bend
16	757	
17	724	
18	557	
19	520	
20	435	
21	414	C–C–C op bend
22	332	Ph–C–Ph

arom: aromatic; bend: bending; str: stretching; ip: in-plane; op: out-of-plane.

stretching vibrational mode of the aromatic ring; the peak at 1363 cm^{-1} is assigned to the *N*-phenyl vibrational mode; the peak at 1172 cm^{-1} is assigned to the C–H in-plane bending vibrational mode; the peaks at 908 and 798 cm^{-1} are assigned to the C–H out-of-plane bending vibrational mode; the peak at 414 cm^{-1} is assigned to the C–C–C out-of-plane bending vibrational mode; and the peak at 332 cm^{-1} is assigned to the phenyl–C–phenyl vibrational mode. These peaks are in good agreement with the results of other authors.[15–17] The peak at 2050 cm^{-1} is a ghost line, which arises from the laser diode that was used as feedback in AFM operation. The exposure time of the liquid-nitrogen-cooled CCD camera was 1sec with 230-μW laser power at the entrance of the inverted microscope. Fig. 8c shows near-field Raman spectrum of crystal violet which is obtained by subtracting the spectrum of Fig. 8b from that of Fig. 8a. Assignment of all Stokes-shifted peaks observed in Fig. 8 is shown in Table 2.

NEAR-FIELD RAMAN IMAGING

Molecular identification at nanometric scale is possible in two dimensions by detecting near-field Raman signal at a specific Stokes line with a monochromator and a point detector while scanning the sample laterally. A metallic tip reduces measurement time for Raman imaging drastically because of the amplification of near-field Raman scattering with locally enhanced field. Consequently, spatial distribution of molecules can be imaged at nanometric scale during reasonable acquisition time.[13] Near-field Raman imaging enables us to attain direct and sensitive observation of many kinds of molecules and their local distributions as well as biological cells without labeling with dye.

In this section, we show such an example of molecular imaging. Mixed aggregates of Rhodamine 6G and crystal violet molecules are used as samples for near-field Raman imaging. The sample is prepared by casting ethanol

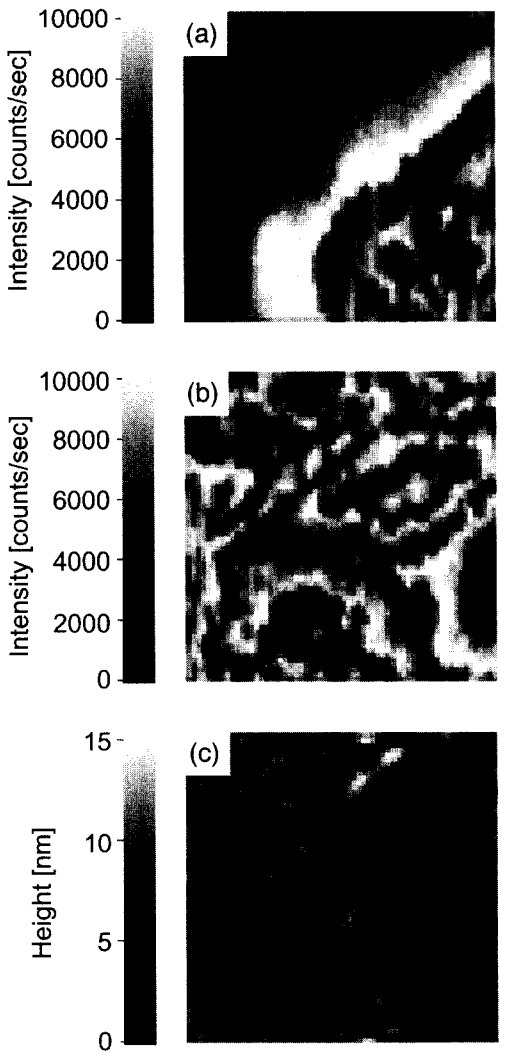


Fig. 9 Near-field Raman images obtained at (a) 607 cm^{-1}; C–C–C in-plane bending mode of Rhodamine 6G and (b) 908 cm^{-1}; C–H out-of-plane bending mode of crystal violet; (c) the corresponding topographic image of the scanned area. It took 10 min to obtain one image where 1×1 μm scanning area consisted of 64×64 pixels.

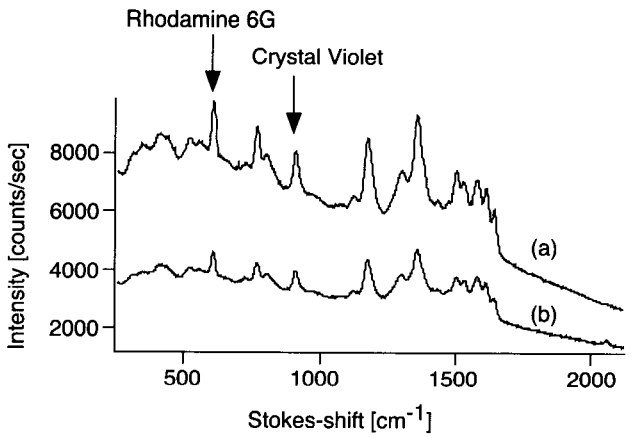


Fig. 10 Typical Raman spectra of the aggregated sample used in Fig. 9, (a) with and (b) without a metallic tip. Stokes-shifted-lines of Rhodamine 6G and crystal violet are observed. Arrows indicate the Stokes-shifted-lines used for near-field Raman imaging in Fig. 9.

solution of Rhodamine 6G and crystal violet on a coverslip coated with an 8-nm-thick silver film and drying the solution. The distribution of molecules for both samples is fairly inhomogeneous and is set to have a 1-nm average thickness of the layer of molecules.

Fig. 9a–c shows near-field Raman images from the same area of the sample, where the concentrations of Rhodamine 6G and crystal violet are 1.25×10^{-3} and 1.25×10^{-2} wt.%, respectively. These concentrations resulted in a comparable Raman scattering intensity from Rhodamine 6G and crystal violet because Rhodamine 6G yields higher scattering cross section than that of crystal violet molecules. Fig. 9a shows an image corresponding to the peak at 607 cm^{-1}, which comes from the Stokes-shifted C–C–C in-plane bending vibrational mode of Rhodamine 6G. Fig. 9b is obtained at 908 cm^{-1} of which peak corresponds to the Stokes-shifted line of the C–H out-of-plane bending vibrational mode of crystal violet. Fig. 9c shows a corresponding topographic image obtained simultaneously with Fig. 9a in AFM operation. The dimension of all the images comprising 64×64 pixels was 1×1 μm. The scanning rate used for each line was 0.1 Hz. At this rate, a single image can be achieved in only 10 min. Fig. 10 shows the Raman spectra of the mixed sample, which exhibits the Stokes-shifted lines of both Rhodamine 6G and crystal violet molecules. The arrows depicted in Fig. 10 point to the Stokes-shifted lines used for the near-field Raman imaging in Fig. 9. We can selectively obtain the distributions of each vibrational mode that we cannot distinguish in the topographic image. The distributions of each vibrational mode are quite different and show complicated structures corresponding to the inhomogeneous distributions of both molecules. According to Fig. 9a, Rhodamine 6G molecules are mainly localized at the lower right position of the figure; on the other hand, Fig. 9b shows that crystal violet molecules are randomly dispersed in the scanned area. In Fig. 9c, the island structures of the silver film are observed in the topographic image because the average thickness of the aggregated molecular layer is estimated to be 1 nm which is much thinner than the silver film (average thickness: 8 nm). Accordingly, the distributions of both molecules are not clearly seen in the topographic image that reflects the pancake structure of the silver grains (30–50 nm in diameter and 8 nm in thickness). Note that without a metallic tip (far-field detection), we could not obtain such high-resolution images because the far-field signal is averaged out within the focused spot [Fig. 2b]. Fig. 9 shows that the near-field Raman images attain the molecular vibrational distributions with a high sensitivity, even if the thickness of the molecular layer is 1 nm. In the experiment, organic dye molecules adsorbed on silver films were used. While the silver films are required for dye

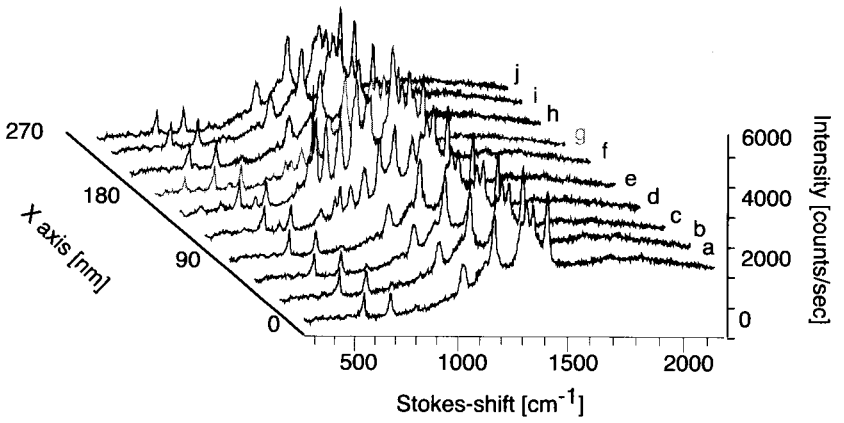


Fig. 11 Near-field Raman spectra mapping at several positions. (*View this art in color at www.dekker.com.*)

molecules to quench the strong fluorescence, near-field Raman spectroscopy and imaging by an apertureless metallic probe scanning optical microscope are applicable not only to organic, but also to inorganic sample even without the support of the silver films. Single molecule detection can be performed after optimizing the tip material, shape,[18] and illumination method[19] because the enhancement factor is very sensitive to the tip and the polarization of the light.

SPECTRAL CHANGE OF NEAR-FIELD RAMAN SPECTRUM

Because near-field Raman spectroscopy using a metallic tip realizes surface-enhanced Raman scattering singly on its apex, chemical enhancement of Raman scattering can be expected besides physically electromagnetic enhancement as mentioned above. Fig. 11 shows a one-dimensional near-field image of the Rhodamine 6G distribution at spatial steps of 30 nm. At each position, the near-field Raman spectrum is detected. The spectra in Fig. 11 are all enhanced by the tip apex, while some of them, e.g., spectra -e, -f, and -g, exhibit remarkable features which are quite different from others or from far-field spectrum. The distinguishable near-field features in those spectra replotted in Fig. 12 are marked with capital alphabetic letters, and the numbers in different spectra correspond to the same assignments as in Fig. 7. Some peaks are slightly shifted and their relative intensities are quite different from those of the far-field spectrum. Furthermore, some additional peaks, which are not observed in the far-field, appear.

For the aromatic C–C stretching vibrational mode, the peaks between 1350 and 1650 cm^{-1} are heavily modified. New shifted peaks, e.g., A: 1457 cm^{-1} in spectrum -e, G: 1483 cm^{-1} in spectrum -f, appeared. For the C–H out-of-plane bending vibrational mode (number 13), peak intensities are constant throughout the scanning, while for the in-plane bending (number 11), weak peaks are strongly enhanced and slightly shifted (see, e.g., J: 1112 cm^{-1}). For the C–O–C bending vibrational mode, the new Stokes-shifted peaks appear and those are largely enhanced (see peaks, e.g., B: 1278 cm^{-1}, H: 1286 cm^{-1}, and M: 1275 cm^{-1}). Furthermore, other new peaks which have not been assigned yet also appear in the spectra (see, e.g., C: 1054 cm^{-1}, D: 1027 cm^{-1}, E: 946 cm^{-1}, F: 902 cm^{-1}, K: 1040 cm^{-1}, L: 705 cm^{-1}, O: 1057 cm^{-1}, P: 1027 cm^{-1}, Q, 700 cm^{-1}, and R: 425 cm^{-1}). Some of the peaks can be identified as the same vibrational modes (e.g., C, K, and O; D and P; L and Q). While some of Stokes-shifted peaks are strongly enhanced, the fluorescence intensity is constant throughout the positions. This

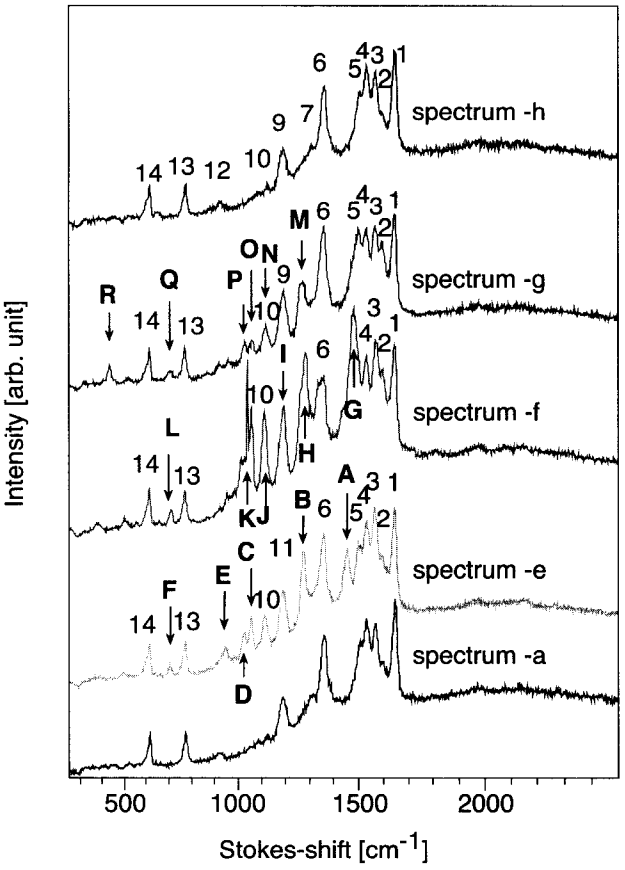


Fig. 12 Some features in the near-field region can be seen and the remarkable near-field characteristic Raman spectra -e, -f, -g in Fig. 11 are plotted. For comparison, spectrum -a and -h are also shown. The numbers correspond to the same assignments as shown in Fig. 7 and remarkable features are pointed out by arrows with capital alphabets (A∼R). (*View this art in color at www.dekker.com.*)

can be explained by the chemical mechanism of surface-enhanced Raman scattering because of the charge transfer excitation between molecules and metal.[20]

CONCLUSION

Application of near-field optics to Raman spectroscopy brings microanalysis of most of materials to their nano-identifications. For example, molecules can be assigned, observed, and detected directly at the nanometric or molecular scale by using this technique without the need for staining a specimen with a dye. Low cross section of Raman scattering is amplified to detectable lever by the locally enhanced field at a metallic tip. Near-field Raman signal is also enhanced by the chemical interaction between molecules and metallic tip. The nanoscale

chemical phenomenon has a potential for determination of molecular orientation. Raman spectroscopy coupled with locally enhanced electromagnetic field, which may contribute to nonlinear spectroscopy such as multiphoton processes, second harmonic generation, and coherent anti-Stokes Raman scattering, will be one of the main topics of microspectroscopic or nanospectroscopic research in the next decade.

REFERENCES

1. *Near-Field Optics and Surface Plasmon Polaritons*; Kawata, S., Ed.; Springer-Verlag: Berlin, Germany, 2001.
2. *Nano-Optics*; Kawata, S., Ohtsu, M., Irie, M., Eds.; Springer-Verlag: Berlin, Germany, 2002.
3. Inouye, Y.; Hayazawa, N.; Hayashi, K.; Sekkat, Z.; Kawata, S. Near-field scanning optical microscope using a metallized cantilever tip for nanospectroscopy. Proc. Soc. Photo-Opt. Instrum. Eng. **1999**, *3791*, 40–48.
4. Hayazawa, N.; Inouye, Y.; Sekkat, Z.; Kawata, S. Metallized tip amplification of near-field Raman scattering. Opt. Commun. **2000**, *183*, 333–336.
5. Bian, R.X.; Dunn, R.C.; Xie, X.S. Single molecule emission characteristics in near-field microscopy. Phys. Rev. Lett. **1995**, *75*, 4772–4775.
6. Inouye, Y.; Kawata, S. Near-field scanning optical microscope using a metallic probe tip. Opt. Lett. **1994**, *19*, 159–161.
7. Zenhausern, F.; O'Boyle, M.P.; Wickramasinghe, H.K. Apertureless near-field optical microscopy. Appl. Phys. Lett. **1994**, *65*, 1623–1625.
8. Bachelot, R.; Gleyzes, P.; Boccara, A.C. Near-field optical microscope based on local perturbation of a diffraction spot. Opt. Lett. **1995**, *20*, 1924–1926.
9. Inouye, Y.; Kawata, S. A scanning near-field optical microscope having scanning electron tunnelling microscope capability using a single metallic probe tip. J. Microsc. **1995**, *178*, 14–19.
10. Furukawa, H.; Kawata, S. Analysis of image formation in a near-field scanning optical microscope: Effects of multiple scattering. Opt. Commun. **1996**, *132*, 170–178.
11. Furukawa, H.; Kawata, S. Local field enhancement with an apertureless near-field-microscope probe. Opt. Commun. **1998**, *148*, 221–224.
12. Hayazawa, N.; Inouye, Y.; Sekkat, Z.; Kawata, S. Near-field Raman scattering enhanced by a metallized tip. Chem. Phys. Lett. **2001**. *335*, 369–374.
13. Hayazawa, N.; Inouye, Y.; Sekkat, Z.; Kawata, S. Near-field Raman imaging of organic molecules by an apertureless metallic probe scanning optical microscope. J. Chem. Phys. **2002**, *117*, 1296–1301.
14. Hayazawa, N.; Inouye, Y.; Kawata, S. Evanescent field excitation and measurement of dye fluorescence using a high N.A. objective lens in a metallic probe near-field scanning optical microscopy. J. Microsc. **1999**, *194*, 472–476.
15. Gicquel, J.; Carles, M.; Bodot, H. Resonance Raman investigation of charge-transfer complexes between a trityl cation (crystal violet) and sulfonated azo derivatives. J. Phys. Chem. **1979**, *83*, 699–706.
16. Sunder, S.; Bernstein, H.J. Resonance Raman-spectrum of a deuterated crystal violet—$[P(CH_3)_2N \cdot C_6D_4]_3C^+Cl^-$. Can. J. Chem. **1981**, *59*, 964–967.
17. Watanabe, T.; Pettinger, B. Surface-enhanced Raman scattering from crystal violet adsorbed on a silver electrode. Chem. Phys. Lett. **1982**, *89*, 501–507.
18. Krug, J.T., II; Sanchez, E.J.; Xie, X.S. Design of near-field optical probes with optimal field enhancement by finite difference time domain electromagnetic simulation. J. Chem. Phys. **2002**, *116*, 10895–10901.
19. Novotny, L.; Sanchez, E.J.; Xie, X. Near-field optical imaging using metal tips illuminated by higher-order Hermite–Gaussian beams. Ultramicroscopy **1998**, *71*, 21–29.
20. Otto, A.; Mrozek, I.; Grabhorn, H.; Akemann, W. Surface-enhanced Raman-scattering. J. Phys. Condens. Matter **1992**, *4*, 1143–1212.

Near-Field Scanning Optical Microscopy: Chemical Imaging

Bogdan Dragnea
Indiana University, Bloomington, Indiana, U.S.A.

INTRODUCTION

Chemical near-field microscopy is a promising optical technique, which transcends the diffraction limit of classical optical microscopy while maintaining its spectroscopic capabilities and sharing the benefit of topographic contrast from proximal probe microscopies. Several techniques and sources of chemical contrast are discussed, from the time-resolved mapping of electric field-induced molecular reorientation in polymer films to the vibrational Raman spectrum of single-carbon nanotubes with 28 nm spatial resolution in the optical image.

BEYOND THE DIFFRACTION LIMIT

With the advent of scanning probe microscopies, the past 20 years or so have witnessed an explosion of analytical tools to explore materials at a nanometer scale. Advances in the understanding of physical and chemical phenomena at these length scales are expected to have a strong impact on microbiology, materials science, and microelectronics. At present, modified viruses are explored as nanobeakers for controlled chemical reactions in confined environments.[1] New materials having exotic optical and electronic properties have been synthesized.[2,3] Novel polymeric resists for sub-100 nm photolithography[4] and alternative schemes for electronic component integration are intensely explored.[5] As a general feature, systems representative of these domains often possess a complex hierarchical organization, which requires the ability to map chemical phenomena on a certain range of length scales.[6] Such studies, more often than not, involve some type of direct imaging technique, which is nonintrusive, has chemical specificity, and has adequate space and time resolution.

Vibrational spectroscopy, encompassing infrared (IR), near-IR, Raman, inelastic neutron scattering, electron energy loss, and cavity ring-down spectroscopy, has exquisite chemical specificity. The positions and intensities of vibrational absorption bands can be used to confirm or identify the presence of a particular group, whereas spectral correlations can be used to access structural and environmental information on selected groups.[7] IR and Raman spectroscopies are particularly nonintrusive, require a small amount of sample, and can be easily coupled with microscopy to provide for spatial resolution.[8,9] The importance of these qualities is reflected in the available range of commercial Raman and Fourier transform infrared (FTIR) microscopes, and in the multitude of applications ranging from materials science to forensics and identification of cultural artifacts. For basic research applications (usually concerning molecular relaxation phenomena), pulsed lasers and fast detectors provide for temporal resolution down to a few femtoseconds.

Classical microscopes achieve a maximum spatial resolution of about 10–20 μm in mid-IR (for wavelengths between 2 and 10 μm), whereas Raman microscopes approach the 1-μm limit. The current spatial resolution limitation in mid-IR is because of the lack of suitable dielectrics to build aberration-corrected lenses.[10] Even if perfect materials were available, there is still a fundamental limitation of imaging instruments based on spherical lenses or mirrors. This fundamental spatial resolution limit is known as the Abbe diffraction limit.[11] Following Abbe's explanation, every object behaves as a superposition of diffraction gratings. In any classical imaging instrument, the front lens is placed, for practical reasons, at least several wavelengths from the object from which the scattered light is collected. The smaller the details of the object are, the higher are the diffraction orders or spatial frequencies of the scattered light. Not all the spatial frequencies will make their way into the lens; there is a cutoff spatial frequency necessarily connected to the geometry of the collection optics. Moreover, the highest spatial frequencies decay exponentially along the object normal.[12] Thus the geometry of the lens and the distance to the sample limit the achievable spatial resolution at: $\Delta x = \lambda/(2\pi \mathrm{NA})$, where λ is the wavelength and NA is the numerical aperture of the lens.

Many problems studied by vibrational microspectroscopy would gain from better spatial resolution. Some of these problems are reviewed later in the section "Applications of Optical Near-Field Microscopy in Chemistry" of this entry. Although the Abbe resolution limit of classical microscopes continues to be pushed in small steps, especially for visible wavelengths, by using special

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009270

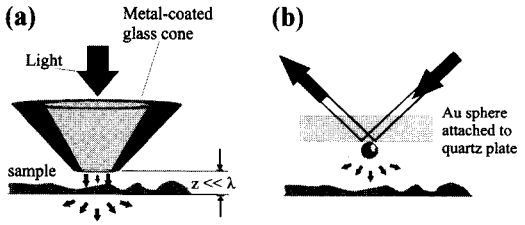


Fig. 1 Synge's proposal for an ultrahigh resolution optical microscope to overcome the diffraction limit. The subwavelength source of light may be (a) a microscopic aperture defined in a metal screen, at the tip of a sharp dielectric cone, and (b) a nanoparticle with large optical polarizability. The proximity of the probe to the surface is an essential element in near-field microscopy.

geometries of illumination[13] or nonlinear optical effects,[14,15] two other conceptually different approaches currently hold the promise of a major breakthrough in ultramicroscopy. The first is the near-field approach, an idea credited to Synge,[16] who proposed the use of a subwavelength source of light such as a tiny hole in a metal screen illuminated by bright light or a gold particle, scattering light from a location very close to the sample surface in a dark-field illumination configuration (Fig. 1). A key element of Synge's approach is the location of the light source very close to the sample surface. As a consequence of the probe proximity, sampling of the evanescent waves bound to the surface occurs. Because in modern instruments based on Synge's idea, the subwavelength light source is scanned across the sample surface by maintaining a constant gap between the surface and the probe to sequentially map the near-field optical interaction, such a microscope is called a near-field scanning optical microscope (NSOM or SNOM),[17] as opposed to far-field instruments, which deal with propagating waves only. The second approach is the negative refraction concept upon which Pendry[18] built his idea of a superlens. Such a lens may overcome the evanescent wave decay and therefore should focus light to considerably smaller areas than that allowed by the Abbe diffraction limit. Although numerous experimental realizations of Synge's proposal have broken the diffraction limit barrier by more than one order of magnitude, the more recent solution based on the Pendry superlens remains in the stage of theoretical concept.

We are concerned here with those techniques and applications of the NSOM principles that have been applied to chemical problems. A number of alternative nonoptical techniques, capable of nanoscale mapping with chemical specificity, will be also mentioned with the goal of situating the near-field optical microspectroscopy within the broader context of chemical mapping methodologies.

APERTURE, APERTURELESS, AND TIPLESS NSOM

Three approaches have been used to date to generate near-field optical images with spatial resolution superior to the diffraction limit. The first approach, and the most often utilized at present, is the one based on a subwavelength optical aperture at the apex of a tapered transparent optical fiber, which has been metal-coated[19] (Fig. 1). The spatial resolution is limited by the size of the aperture and the electromagnetic skin depth of the metal coating. The second method is based on an apertureless sharp tip, usually metallic, which is acting as a scattering probe, converting surface sample-bound evanescent waves into propagating waves (Fig. 2). The spatial resolution attainable with this method is determined by the tip radius, rather than the electromagnetic skin depth of the metal.[20] The third proposal does not use a tip. The probe in this case can be a nanoparticle, which is manipulated by optical tweezers to scan the sample surface.[21] The theory of this microscope, including the excitation of particle eigenmodes, has been worked out by Lester et al.[22] Tischer et al. have demonstrated a related nanoparticle-based technique that is able to provide topological three-dimensional (3-D) images of nanoscopic cavities accessible to the particle with 20-nm spatial resolution normal to the immersed surfaces and ~200-nm lateral resolution. The accessible volumes are reconstructed from the histogram of thermal position fluctuations of the particle, hence the name of the technique—3-D thermal noise imaging.[23] The 3-D thermal noise imaging has not been used yet in conjunction with near-field measurements. Palanker et al.[24] have used a localized nonequilibrium plasma concentration on a semiconductor surface photoinduced by a tightly focused laser beam operating at visible wavelengths as a scanning local probe for near-field IR reflectivity measurements.

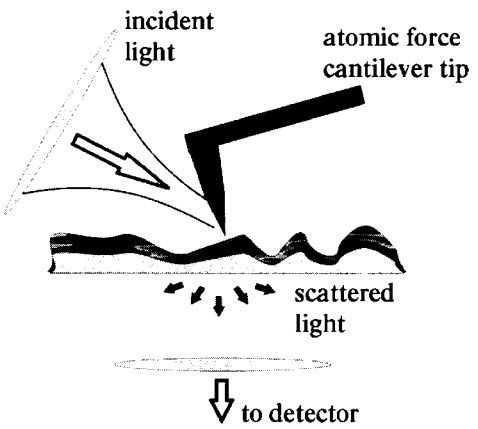


Fig. 2 Schematic of the optical probe region in an apertureless NSOM with grazing illumination.

Among the three near-field optical schemes, only the optical fiber NSOM and the apertureless NSOM have been applied to the spatial mapping of chemical properties. Nevertheless, the nanoparticle approach holds a special appeal because of intrinsic field enhancements at the particle surface,[25] which may provide for 3-D surface-enhanced Raman or IR microspectroscopy of fluid-immersed interfaces in the future.[26] Several comprehensive reviews on the theory and applications of the optical fiber NSOM and the apertureless NSOM, especially in the visible range of the spectrum, have been published to date.[27–30] Particularities and chemical applications of the mid-IR NSOM have been reviewed by Dragnea and Leone[31] and Keilman.[32] The far-IR and the microwave region have been covered by Rosner and van der Weide.[33]

Choosing an apertureless or a fiber optic aperture NSOM for chemical characterization depends on the type of application. The fiber-based instrument is closer to a scanning confocal microscope, which, instead of the NSOM aperture, would use the real image of a pinhole formed through a lens system. Therefore contrast formation is arguably more intuitive. An important advantage of the aperture scheme is negligible background. For a perfect metal coating, the only light comes from the subwavelength aperture. As a consequence, sensitive detection schemes such as photon counting can be used. Bringing the light close to the sample through a fiber may represent a very important feature, which makes the aperture NSOM unique for certain applications dealing with samples immersed in an absorbing fluid. For example, Hong et al.[34] have succeeded to image living cells in water using an IR near-field microscope. Such studies are not possible in far field or with an apertureless NSOM because of the strong attenuation of the incident beam caused by water absorption bands associated with H–O–H bending, which overlap with the protein amide bands at ~ 1650 cm^{-1}. However, in aperture NSOM, because the gap between the fiber probe aperture and the cell membrane (sample surface) is of a few nanometers, most of the signal comes from absorption by the cellular protein and not by the water layers between the probe and the sample surface. A disadvantage of the optical fiber approach is low transmission (10^{-4}–10^{-6}, depending on the wavelength and on the fiber taper profile), combined with a limit on the incident light power dictated by the damage threshold of the metal coating (~ 10 mW for Al-coated fibers). The throughput problem is a challenge in particular at longer wavelengths because the cutoff diameter for guided propagation occurs earlier along the taper than for shorter wavelengths, and the light has to propagate evanescently over longer distances until it reaches the aperture. At IR wavelengths, one usually has to reach a compromise between spatial resolution, determined by the aperture diameter, and sensitivity. The fiber probe fabrication remains, at present, the main challenge and a key element for further improvements of the aperture NSOM, at least at IR wavelengths.

The apertureless NSOM has four advantages on its side, with respect to the aperture NSOM. 1) The attainable optical resolution exceeds that of the aperture-based microscopes. This is because of the fact that, for a suitable incident field polarization, field enhancement occurs in a very localized area of the size of the tip radius.[35] 2) The problem of the strong attenuation in the fiber taper is alleviated in the apertureless scheme because the light is focused from the far field directly on the tip, without loss. This is why the apertureless NSOM has, until now, dominated microspectroscopy applications in the mid-IR and far-IR ranges, where the losses within the fiber taper are the largest. 3) The tip fabrication step can be avoided altogether because a large selection of commercial atomic force microscopy (AFM) tips is available. Apertureless tips are more robust and their fabrication is more reproducible than that of coated optical fiber tips. 4) The range of usable wavelengths is not limited, such as in the case of fiber-based NSOMs by material properties. The main disadvantage of the apertureless NSOM comes from strong scattered light in areas of the tip shaft, far from the sample, which creates a strong background for the detector. Furthermore, the background usually varies during scanning because of the z-motion of the tip in constant force scanning mode, hence the strong possibility of topographic artifacts. Great care has to be taken to decouple the topographic artifacts, which often adds complexity to the experimental setup. However, several strategies have been found to eliminate this problem from apertureless NSOM images.[20,36]

Whether an apertureless or a fiber optic aperture NSOM is chosen, two issues have to be considered while interpreting the data: one is the mechanism of contrast generation; the other is scanning probe imaging artifacts.[37] The electromagnetic field distribution around a nanoscopic object is extremely sensitive to both physical characteristics of the object and the illumination mode. A rigorous description of the interaction optical between the probe and the field close to the surface is necessary to properly describe the image formation. Numerous simulation schemes for numerically solving Maxwell's equations have been used. The practical difference between these schemes lies in how the challenge of sharp corners and edges, much smaller than the wavelength, is handled.[38] Palanker et al.[39] have provided an approximate phenomenological model for contrast formation in aperture IR-NSOM. The model relies on a photon tunneling picture in which the size of the barrier representing the tip/gap/surface system is determined experimentally

from tip–surface approach curves. The model correctly reproduces the observed differences between near-field and far-field IR absorption spectra obtained experimentally (Fig. 3). An analytical approximation for the scattered field in the apertureless case has been proposed by Hillenbrand et al.[36]

To avoid z-motion crosstalk in the optical images, the safest way is to turn off the force feedback on the z-piezo and scan at constant height. However, if constant gap scanning is required to be certain that the near-field optical images are not merely topographic artifacts, according Hecht et al.,[37] at least one of the following conditions should be satisfied:

a) Near-field images and topographic images are uncorrelated. This means that often the best way to use a near-field microscope is on samples that are flat, but have heterogeneities with strong optical contrast.

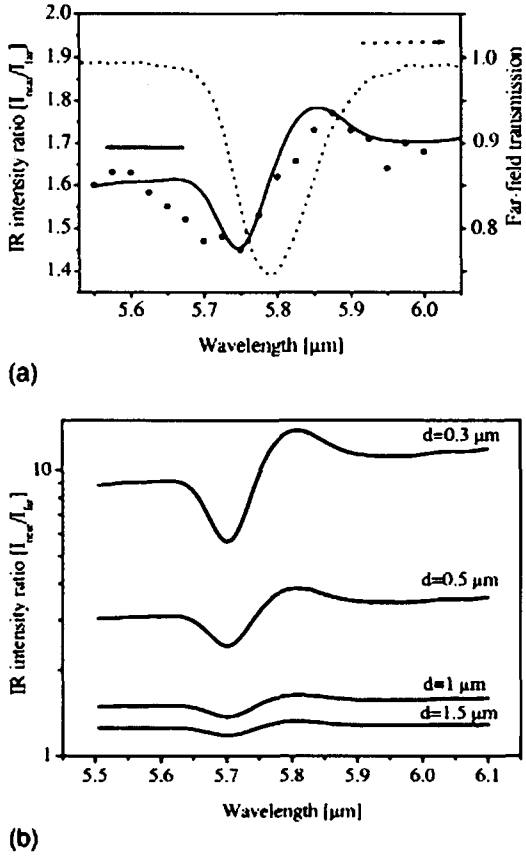


Fig. 3 (a) Differences between experimental infrared near-field (symbols) and far-field (dashed line) spectra. The continuous line represents a fit generated from a phenomenological model based on the photon tunneling picture. (b) Calculated spectra for tip diameters ranging from 0.3 to 1.5 μm. The amplitude of the near-field modulation is inversely proportional to the aperture diameter. (From Ref. [39], with permission)

b) When correlations between the topographic image and the optical image are apparent, they should be displaced by a constant amount.

c) The spatial resolution of the optical image is different from that of the topographic image.

APPLICATIONS OF OPTICAL NEAR-FIELD MICROSCOPY IN CHEMISTRY

In a few NSOM applications, which take advantage of the presence of narrow absorption bands of the sample and tunable light sources, such as in the case of vibrational microspectroscopy or the fluorescence emission properties of the sample, the near-field image contrast formation can be at least qualitatively understood in a more straightforward way than by numerically solving the Maxwell equations. A few examples from this category are described below.

Fluorescence NSOM

In fluorescence NSOM, the chemical contrast is achieved by measuring the local fluorescence spectroscopy of the sample. Because the fluorescence wavelengths of many organic materials lie in the visible range, this method has been extensively applied to organic materials. The visible wavelength range has the benefit of sensitive detectors, optimized optics, and powerful and relatively low-cost laser sources. These features, plus the fact that the strong excitation background can be easily filtered out, made fluorescence NSOM reach single-molecule sensitivity at room temperature.[40] Fluorescence NSOM has been used to directly identify molecular aggregations in amorphous conjugated polymer films of approximately 100 nm, which were detrimental to the photoluminescent and electronic properties of the film.[41] Such studies have clearly shown that the NSOM technique can contribute to a better understanding of the technologically important thin organic films. Barbara et al.[42] have reviewed the characterization of complex organic thin film materials by time-resolved and wavelength-resolved near-field microspectroscopy. In organic materials for electronic devices, the device fabrication processes induce morphological changes at length scales for which NSOM became the method of choice. More recently, fluorescence NSOM investigations have proven instrumental in the direct observation of the influence of device operating conditions on local organic material morphology.[43] Fig. 4 is an illustration of how the dynamics of field-induced local reorientation of polymer-dispersed liquid crystal films can be measured using fluorescence NSOM.

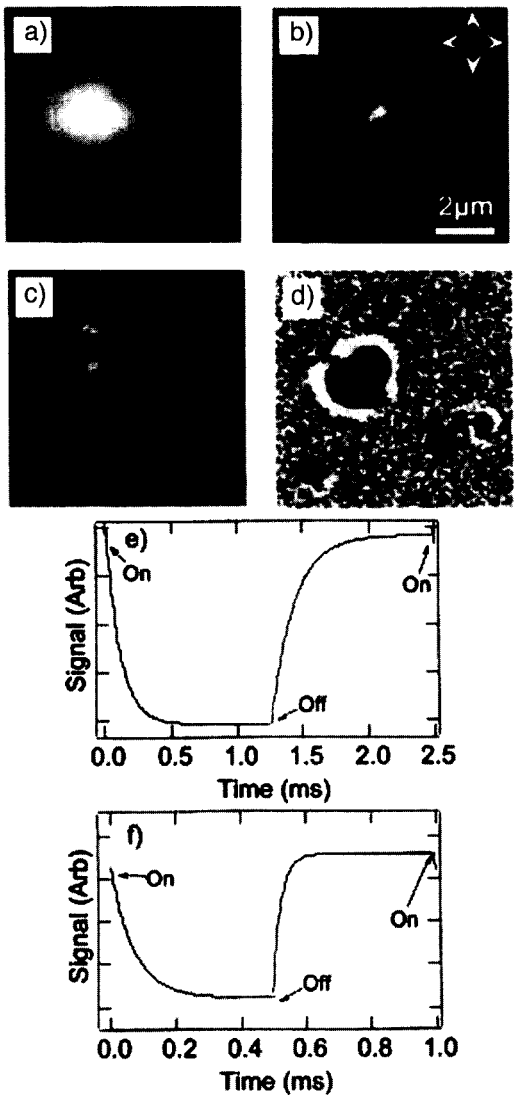


Fig. 4 Measuring the local field-induced reorientation in molecular films. A sinusoidal electric bias is applied between the Al coating of the fiber probe and the transparent support electrode on which a polymer-dispersed liquid crystal film is deposited. Frame (a) represents the shear-force topographic image acquired simultaneously with near-field optical images (b–d). Frames represent the modulated optical signal as a function of the applied voltage. (d) A map of the phase lag of the optical modulation signal with respect to the applied bias. Frames (e) and (f) represent time-dependent variations in the transmitted optical signal from central and edge probe locations. The relaxation rate is faster in the edge region. (From Ref. [42].)

Raman Near-Field Microscopy

When chemical identification is required, few methods can be as specific as vibrational spectroscopy. In vibrational Raman spectroscopy, the chemical contrast comes from a nonlinear effect: the molecule polarizability is modulated, depending on the molecular symmetry, by vibrational interactions. Because of its nonlinear origin, the Raman scattering has a much smaller molecular cross section (10^{-31}–10^{-29} cm^2) when compared with typical IR absorption cross sections (10^{-17}–10^{-16} cm^2). However, when a molecule is placed in close vicinity to a rough metal surface, the Raman scattering may be enhanced by an average factor of 10^6. The surface enhancement is thought to be of both chemical and electromagnetic origin; however, the separation of contributions from these two effects continues to be elusive.[25] An additional enhancement for surface-enhanced Raman scattering (SERS) of approximately 10^5 occurs when there is resonance between the pump wavelength and a molecular electronic transition. Even when combining these enhancements, the resulting cross section seems to be too small to be used to generate aperture NSOM contrast, if one takes into consideration the limit imposed by fiber damage threshold on the pump power, ~10 mW. However, Emory and Nie[44] and Kneipp et al.[45] showed that the aforementioned enhancements represent only ensemble-averaged values, and that selected silver colloidal particles may exhibit individual Raman enhancement factors of 10^{14}–10^{15}, or effective molecular cross sections for Raman scattering comparable with linear IR absorption. The discovery boosted the interest in the possibility of using the surface enhancement to build near-field Raman scattering probes. The apertureless NSOM approach has been the method of choice in this endeavor. In this case, the nanostructure generating SERS enhancement is the tip itself. To establish a strong electromagnetic local enhancement, the electric field polarization of the pump laser beam has to be oriented along the tip axis.[35] Two experimental apertureless schemes have been implemented, different in respect to the way in which this condition is accomplished (Fig. 5). The first approach has been to use the side illumination of the tip.[46] Its advantage is that opaque samples can be

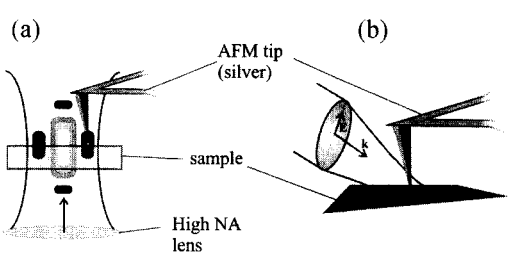


Fig. 5 Apertureless Raman NSOM approaches. (a) On-axis illumination. The contour lines in the center of the focused beam represent schematically the distribution of regions of equal intensity. (b) Grazing incidence illumination.

imaged. The disadvantage is that long working distance Raman objectives have to be used, which limits the collection efficiency. An alternative illumination technique is to use on-axis illumination by strongly focused light, with the tip displaced from the center of the beam in the direction of polarization in one of the side lobes of the focal region.[47] An impressive optical 25-nm spatial resolution and a photon rate of a few thousands of photons per second have been obtained on isolated single-walled carbon nanotubes using this approach by Hartschuh et al. Because of the large Raman signal, acquisition of simultaneous near-field Raman and topographic images is possible. In this work, local variations in the Raman spectrum along a carbon nanotube have been found (Fig. 6). A clear proof of the near-field origin of the Raman enhancement has been made through tip–sample distance measurements.[47] The Raman spectrum local variations reflect changes in the molecular structure, which can be caused by defects, the presence of catalyst particles, or variations in the tube lattice, clearly demonstrating the promise that SERS–NSOM hold for experimental investigations of individual nanostructures, perhaps down to the single macromolecule level.

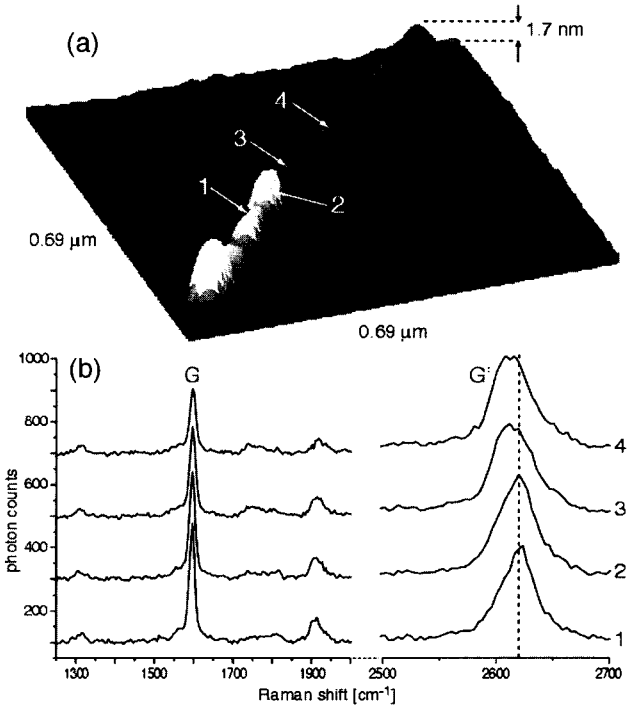


Fig. 6 (a) Topographic image of a single-wall nanotube on glass. (b) Near-field Raman spectra detected at the marked positions 1–4 in (a). The three bumps in the topographic image are probably metal catalyst nanoparticles. (From Ref. [47].)

Infrared Near-Field Microscopy

The impressive chemical sensitivity and spatial resolution of the Raman NSOM are based on the local enhancement of the field very close to the tip. The decrease in enhancement is predicted to vary as $(a+H)^{-12}$, with the distance H from the tip, where a is the radius of the tip.[25] This means that apertureless Raman SERS is essentially a surface technique, which probes a 1- to 2-nm-thick layer. In the case of aperture NSOM, the axial length of the probing region roughly corresponds to the diameter of the aperture. For a 100-nm aperture, one may probe a $\sim$100-nm film without loss in lateral resolution. However, the aperture scheme has not made the same progress as the apertureless NSOM in Raman microscopy because of the absence of surface enhancement. An alternative is IR near-field microscopy.

Mapping the chemical composition of thin films by transmission IR-NSOM has proven particularly useful in the characterization of polymeric resists for deep ultraviolet (DUV) nanolithography.[48] In this example, IR absorption near-field spectroscopy has been employed by taking advantage of the "fingerprint" region of the molecular spectra of different compounds resulting from the DUV-patterned exposure of the resist. The goal of this work was the characterization of the latent image formation during the technological steps of proximity mask-assisted DUV patterning and postexposure bake. Two polymeric species were present: poly(methacrylic acid) in the exposed regions and poly(*tert*-butylmethacrylate) in the unexposed regions, which is a photoresist from the acrylate family showing promise for DUV lithography. The most prominent differences in the IR spectra of the two polymers reside in the 2500–3500 cm^{-1} region, because of the hydroxyl absorption of the poly(methacrylic acid), and in the 800–1500 cm^{-1} region, because of the loss of the *t*-butoxy group (Fig. 7a). Tuning the wavelength of the color center laser source on and off, the OH stretch absorption band resulted in generating chemical contrast in near-field images of the latent pattern (Fig. 7b). The attained optical resolution was approximately 280 nm for a 3-μm wavelength. McDonough et al.[49] have extended this work in a promising direction by measuring the water vapor uptake in photolithographic polymers in a controlled environment.

Although the IR fingerprint region offers a very convenient way for chemical identification, there are three main experimental challenges to consider when working with IR-NSOM. The first is the limited availability of tunable laser sources in IR. The closest to the ideal source from the point of view of spectral brightness, tunability, and collimation is the free electron laser,[50] which has the best chances at present to support the realization of an IR near-field microspectroscopic instrument

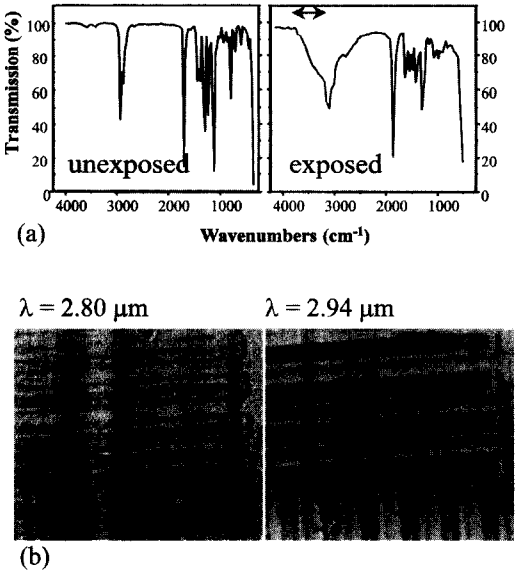


Fig. 7 (a) Infrared spectra of as-coated (left) and irradiated and postexposure baked polymeric resist films (right). The horizontal double arrow represents the tunable wavelength range of the IR color center laser. (b) IR-NSOM mapping of 8 μm/8 μm line/space pattern latent images written by DUV exposure on a 1-μm-thick chemically amplified photoresist film. The images are taken at wavelengths outside (2.80 μm) and within (2.94 μm) the OH stretch absorption band of the exposed polymer, in constant height mode (shear–force coupling with the sample disabled).

that would surpass the spatial resolution of an FTIR microscope, while covering the same spectral range. Another possible approach is to use a tabletop ultrafast broadband laser source. By difference frequency mixing, the signal and the idler beams in a nonlinear crystal tunable IR radiation between 2.5 and 12 μm can be obtained at reasonable average powers (around 10 mW).[51] Another popular choice, especially for apertureless IR-NSOMs, are CO_2 lasers. Knoll and Keilmann[52] used a tunable CO_2 laser coupled with such a microscope to demonstrate vibrational band contrast around $\lambda = 10$ μm on polystyrene particles embedded in polymethyl methacrylate. On the biological applications side, Akhremitchev et al. reported on monolayer-sensitive imaging of DNA hexadecanethiol stripes, using a similar instrument. In this case, chemical contrast was achieved because of absorption by phosphate stretching band of DNA.[53] Interestingly, the measured infrared contrast, free of topographic artifacts, was significantly greater than the calculated one in both examples. This result was interpreted as evidence for surface-enhanced IR absorption.[54]

Besides the light source challenge, other experimental issues to be aware of when dealing with IR-NSOM are: probe optimization, to increase coupling with the near-field; and improved detection, because the detectivity of available detectors for IR is approximately four orders of magnitude less than in the visible. Moreover, vibrational absorptions in IR are a few orders of magnitude weaker than the average electronic absorptions in the visible. These aspects have been considered in detail by Dragnea and Leone[31] in their review of submicron infrared vibrational band chemical imaging.

Nonlinear Near-Field Spectroscopy

The various field enhancements present in the near field make possible the observation and use of coherent nonlinear effects for near-field imaging. Nonlinear techniques are appealing because: 1) the signal wavelength is shifted with respect to the pump wavelength. As a consequence, these techniques have low background and, sometimes, the signal wavelength is situated in a spectral region where detectors have higher sensitivity (such as in the IR–visible sum frequency generation, a vibrational spectroscopy where the measurements are done in the visible although molecular vibrational bands are probed). 2) Based on same principles as SERS, resonant enhancements may boost the nonlinear optical response by orders of magnitude. It follows that in certain situations, the nonlinear optical contrast will be significantly enhanced with respect to linear absorption microspectroscopy. 3) Within the dipole approximation, depending on the symmetry of the sample and the order of the nonlinear effect in use, interfacial or bulk selectivity can be obtained. Techniques and applications of nonlinear chemical imaging microspectroscopy have been reviewed by Schaller et al.[55]

OTHER TECHNIQUES OF CHEMICAL MICROSCOPY

The resolution range of near-field optical microscopy covers length scales between a single molecule and a few microns. At the smallest scales, it overlaps with scanning tunneling microscopy (STM), from the same family of scanning probe techniques. The single-molecule chemistry, imaging, and manipulation by STM have been described in detail in a recent review by Ho.[56] An alternative scanning probe technique with chemical specificity is chemical force microscopy (CFM), a variant of AFM.[57] CFM overcomes the chemical nonspecificity of AFM by covering the tips with well-defined layers of molecular groups having specific interactions with the surface species to be studied.[57] CFM is a complementary technique to chemical NSOM in several respects: first, it is a surface technique, which probes only the topmost

atomic layers of the sample; second, chemical identification of unknown compounds is not possible—a priori knowledge of the sample is necessary; third, transient species cannot be studied by CFM, which is essentially a static method; and, fourth, the contrast mechanism arises from different molecular forces between the tip and the surface, whereas in optical microscopy, chemical information is usually gathered using molecular vibrational or electronic transitions.

CONCLUSION

Instrumental advances in chemical near-field optical microscopy have been the focus of the majority of articles published in the past decade in this area. However, at present, the field starts to emerge from the stage of method development. The selected examples mentioned above have been chosen to provide an illustration of notable concepts, which already have begun to shape the research areas where they have been applied, rather than to faithfully illustrate the breath of this rapidly emerging field. At present, these applications mostly span topics in molecular materials, where interesting property variations occur at submicron length scales. Although the potential of the chemical near-field microscopy for biological sciences has been demonstrated with high-resolution Raman, IR near-field, and nonlinear optical spectroscopy, significant contributions to solving problems in this area are still to come.

REFERENCES

1. Douglas, T.; Young, M. Host–guest encapsulation of materials by assembled virus protein cages. Nature **1998**, *393* (6681), 152–155.
2. Alivisatos, A.P.; Barbara, P.F.; Castleman, A.W.; Chang, J.; Dixon, D.A.; Klein, M.L.; McLendon, G.L.; Miller, J.S.; Ratner, M.A.; Rossky, P.J.; Stupp, S.I.; Thompson, M.E. From molecules to materials: Current trends and future directions. Adv. Mater. **1998**, *10* (16), 1297–1336.
3. Markovich, G.; Collier, C.P.; Henrichs, S.E.; Remacle, F.; Levine, R.D.; Heath, J.R. Architectonic quantum dot solids. Acc. Chem. Res. **1999**, *32* (5), 415–423.
4. Wallraff, G.M.; Hinsberg, W.D. Lithographic imaging techniques for the formation of nanoscopic features. Chem. Rev. **1999**, *99* (3), 1801–1821.
5. Nealey, P.F.; Black, A.J.; Wilbur, J.L.; Whitesides, G.M. Micro- and Nanofabrication Techniques Based on Self-Assembled Monolayers. In *Molecular Electronics*; Jortner, J., Ratner, M., Eds.; Blackwell Science: Osney Mead, Oxford, 1997.
6. Whitesides, G.M.; Boncheva, M. Beyond molecules: Self-assembly of mesoscopic and macroscopic components. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99* (8), 4769–4774.
7. Chalmers, J.M.; Griffiths, P.R. *Handbook of Vibrational Spectroscopy*; John Wiley and Sons: New York, 2002.
8. Wetzel, D.L.; LeVine, S.M. Microspectroscopy—Imaging molecular chemistry with infrared microscopy. Science **1999**, *285* (5431), 1224–1225.
9. Fujii, G. To fuse or not to fuse: The effects of electrostatic interactions, hydrophobic forces, and structural amphiphilicity on protein-mediated membrane destabilization. Adv. Drug Deliv. Rev. **1999**, *38* (3), 257–277.
10. Humecki, H.J. *Practical Guide to Infrared Microspectroscopy*; Marcel Dekker: New York, 1995.
11. Born, M.; Wolf, E. *Principles of Optics*, 6th Ed.; Cambridge University Press: Cambridge, 1980.
12. Goodman, J.W. *Introduction to Fourier Optics, Physical and Quantum Electronics*; McGraw-Hill: New York, 1968.
13. Schrader, M.; Bahlmann, K.; Giese, G.; Hell, S.W. 4Pi-confocal imaging in fixed biological specimens. Biophys. J. **1998**, *75* (4), 1659–1668.
14. Klar, T.A.; Jakobs, S.; Dyba, M.; Egner, A.; Hell, S.W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. Proc. Natl. Acad. Sci. U. S. A. **2000**, *97* (15), 8206–8210.
15. Williams, R.M.; Zipfel, W.R.; Webb, W.W. Multiphoton microscopy in biological research. Curr. Opin. Chem. Biol. **2001**, *5* (5), 603–608.
16. Synge, E.H. Suggested method for extending microscopic resolution into the ultra-microscopic region. Philos. Mag. **1928**, *6* (2), 356–362.
17. Courjon, D. *Near-Field Microscopy and Near-Field Optics, 1*; Imperial College Press: London, 2003.
18. Pendry, J.B. Negative refraction makes a perfect lens. Phys. Rev. Lett. **2000**, *85* (18), 3966–3969.
19. Betzig, E.; Isaacson, M.; Lewis, A. Collection mode near-field scanning optical microscopy. Appl. Phys. Lett. **1987**, *51* (25), 2088–2090.
20. Zenhausern, F.; Martin, Y.; Wickramasinghe, H.K. Scanning interferometric apertureless microscopy—Optical imaging at 10 Angstrom resolution. Science **1995**, *269* (5227), 1083–1085.
21. Sugiura, T.; Okada, T.; Inouye, Y.; Nakamura, O.; Kawata, S. Gold-bead scanning near-field optical microscope with laser-force position control. Opt. Lett. **1997**, *22* (22), 1663–1665.
22. Lester, M.; Arias-Gonzalez, J.R.; Nieto-Vesperinas,

M. Fundamentals and model of photonic-force microscopy. Opt. Lett. **2001**, *26* (10), 707–709.
23. Tischer, C.; Altmann, S.; Fisinger, S.; Horber, J.K.H.; Stelzer, E.H.K.; Florin, E.L. Three-dimensional thermal noise imaging. Appl. Phys. Lett. **2001**, *79* (23), 3878–3880.
24. Palanker, D.V.; Knippels, G.M.H.; Smith, T.I.; Schwettman, H.A. IR microscopy with a transient photo-induced near-field probe (tipless near-field microscopy). Opt. Commun. **1998**, *148* (4–6), 215–220.
25. Moskovits, M. Surface–enhanced spectroscopy. Rev. Mod. Phys. **1985**, *57* (3), 783–826.
26. Kneipp, K.; Kneipp, H.; Itzkan, I.; Dasari, R.R.; Feld, M.S. Surface-enhanced Raman scattering and biophysics. J. Phys., Condens. Matter **2002**, *14* (18), R597–R624.
27. Pohl, D.W.; Bach, H.; Bopp, M.A.; Deckert, V.; Descouts, P.; Eckert, R.; Guntherodt, H.J.; Hafner, C.; Hecht, B.; Heinzelmann, H.; Huser, T.; Jobin, M.; Keller, U.; Lacoste, T.; Lambelet, P.; Marquis-Weible, F.; Martin, O.J.F.; Meixner, A.J.; Nechay, B.; Novotny, L.; Pfeiffer, M.; Philipona, C.; Plakhotnik, T.; Renn, A.; Sayah, A.; Segura, J.M.; Sick, B.; Siegner, U.; Tarrach, G.; Vahldieck, R.; Wild, U.P.; Zeisel, D.; Zenobi, R. Optical microscopy in the nano-world. Chimia **1997**, *51* (10), 760–767.
28. Hecht, B.; Sick, B.; Wild, U.P.; Deckert, V.; Zenobi, R.; Martin, O.J.F.; Pohl, D.W. Scanning near-field optical microscopy with aperture probes: Fundamentals and applications. J. Chem. Phys. **2000**, *112* (18), 7761–7774.
29. Dunn, R.C. Near-field scanning optical microscopy. Chem. Rev. **1999**, *99* (10), 2891.
30. Dereux, A.; Girard, C.; Weeber, J.C. Theoretical principles of near-field optical microscopies and spectroscopies. J. Chem. Phys. **2000**, *112* (18), 7775–7789.
31. Dragnea, B.; Leone, S.R. Advances in submicron infrared vibrational band chemical imaging. Int. Rev. Phys. Chem. **2001**, *20* (1), 59–92.
32. Keilman, F. Vibrational-infrared near-field microscopy. Vib. Spectrosc. **2002**, *29* (1–2), 109–114.
33. Rosner, B.T.; van der Weide, D.W. High-frequency near-field microscopy. Rev. Sci. Instrum. **2002**, *73* (7), 2505–2525.
34. Hong, M.K.; Jeung, A.G.; Dokholyan, N.V.; Smith, T.I.; Schwettman, H.A.; Huie, P.; Erramilli, S. Imaging single living cells with a scanning near-field infrared microscope based on a free electron laser. Nucl. Instrum. Methods Phys. Res., B Beam Interact. Mater. Atoms **1998**, *144* (1–4), 246–255.
35. Wessel, J. Surface-enhanced optical microscopy. J. Opt. Soc. Am., B, Opt. Phys. **1985**, *2* (9), 1538–1541.
36. Hillenbrand, R.; Knoll, B.; Keilmann, F. Pure optical contrast in scattering-type scanning near-field microscopy. J. Microsc. Oxf. **2001**, *202* (77–83).
37. Hecht, B.; Bielefeldt, H.; Inouye, Y.; Pohl, D.W.; Novotny, L. Facts and artifacts in near-field optical microscopy. J. Appl. Phys. **1997**, *81* (6), 2492–2498.
38. Girard, C.; Joachim, C.; Gauthier, S. The physics of near-field. Rep. Prog. Phys. **2000**, *63* (893–938).
39. Palanker, D.V.; Simanovskii, D.M.; Huie, P.; Smith, T.I. On contrast parameters and topographic artifacts in near-field infrared microscopy. J. Appl. Phys. **2000**, *88* (11), 6808–6814.
40. Trautman, J.K.; Macklin, J.J.; Brus, L.E.; Betzig, E. Near-field spectroscopy of single molecules at room-temperature. Nature **1994**, *369* (6475), 40–42.
41. DeAro, J.A.; Weston, K.D.; Buratto, S.K.; Lemmer, U. Mesoscale optical properties of conjugated polymers probed by near-field scanning optical microscopy. Chem. Phys. Lett. **1997**, *277* (5–6), 532–538.
42. Barbara, P.F.; Adams, D.M.; O'Connor, D.B. Characterization of organic thin film materials with near-field scanning optical microscopy (NSOM). Annu. Rev. Mater. Sci. **1999**, *29* (433).
43. McNeill, J.D.; O'Connor, D.B.; Barbara, P.F. Imaging organic device function with near-field scanning optical microscopy. J. Chem. Phys. **2000**, *112* (18), 7811–7821.
44. Emory, S.R.; Nie, S.M. Near-field surface-enhanced Raman spectroscopy on single silver nanoparticles. Anal. Chem. **1997**, *69* (14), 2631–2635.
45. Kneipp, K.; Wang, Y.; Kneipp, H.; Perelman, L.T.; Itzkan, I.; Dasari, R.; Feld, M.S. Single molecule detection using surface-enhanced Raman scattering (SERS). Phys. Rev. Lett. **1997**, *78* (9), 1667–1670.
46. Anderson, M.S.; Pike, W.T. A Raman-atomic force microscope for apertureless—Near-field spectroscopy and optical trapping. Rev. Sci. Instrum. **2002**, *73* (3), 1198–1203.
47. Hartschuh, A.; Anderson, N.; Novotny, L. Near-field Raman spectroscopy using a sharp metal tip. J. Microsc. Oxf. **2003**, *210* (234–240).
48. Dragnea, B.; Preusser, J.; Szarko, J.M.; Leone, S.R.; Hinsberg, W.D. Pattern characterization of deep-ultraviolet photoresists by near-field infrared microscopy. J. Vac. Sci. Technol., B **2001**, *19* (1), 142–152.
49. McDonough, L.A.; Dragnea, B.; Preusser, J.; Leone, S.R.; Hinsberg, W.D. Water vapor uptake in

photolithographic polymers observed by infrared near-field scanning optical microscopy in a controlled environment. J. Phys. Chem., B **2003**, *107* (21), 4951–4954.

50. Cricenti, A.; Generosi, R.; Luce, M.; Perfetti, P.; Margaritondo, G.; Talley, D.; Sanghera, J.S.; Aggarwal, I.D.; Tolk, N.H. Very high resolution near-field chemical imaging using an infrared free electron laser. Phys. Chem. Chem. Phys. **2002**, *4* (12), 2738–2741.
51. Michaels, C.A.; Stranick, S.J.; Richter, L.J.; Cavanagh, R.R. Scanning near-field infrared microscopy and spectroscopy with a broadband laser source. J. Appl. Phys. **2000**, *88* (8), 4832–4839.
52. Knoll, B.; Keilmann, F. Near-field probing of vibrational absorption for chemical microscopy. Nature **1999**, *399* (6732), 134–137.
53. Akhremitchev, B.B.; Sun, Y.J.; Stebounova, L.; Walker, G.C. Monolayer-sensitive infrared imaging of DNA stripes using apertureless near-field microscopy. Langmuir **2002**, *18* (14), 5325–5328.
54. Osawa, M. Surface-Enhanced Infrared Absorption. In *81 Near-Field Optics and Surface Plasmon Polaritons*; Springer-Verlag, 2001.
55. Schaller, R.D.; Johnson, J.C.; Wilson, K.R.; Lee, L.F.; Haber, L.H.; Saykally, R.J. Nonlinear chemical imaging nanomicroscopy: From second and third harmonic generation to multiplex (broadbandwidth) sum frequency generation near-field scanning optical microscopy. J. Phys. Chem., B **2002**, *106* (20), 5143–5154.
56. Ho, W. Single-molecule chemistry. J. Chem. Phys. **2002**, *117* (24), 11033–11061.
57. Noy, A.; Sanders, C.H.; Vezenov, D.V.; Wong, S.S.; Lieber, C.M. Chemically-sensitive imaging in tapping mode by chemical force microscopy: Relationship between phase lag and adhesion. Langmuir **1998**, *14* (7), 1508–1511.

Nucleation of Nanoparticles in Ultrathin Polymer Films

Pieter Stroeve
University of California, Davis, California, U.S.A.

INTRODUCTION

In recent years, the fabrication of nanostructured materials and exploration of their properties have attracted the attention of physicists, chemists, biologists, and engineers.[1–4] Interest in such systems arises from the fact that the mechanical, chemical, electrical, optical, magnetic, electrooptical, and magnetooptical properties of these particles are different from their bulk properties and depend on the particle size.[5–9] There are numerous areas where nanoparticulate systems are of scientific and technological interest. To produce these systems in appropriate matrices is a significant challenge. This article considers nanoparticles in nanofilms of polyelectrolyte. First, we will briefly review some key areas of applications of nanoparticle systems, with an emphasis on nanoparticles in matrices, and then we will address the area of layer-by-layer deposition of polyelectrolytes. Finally, nucleation and growth of nanoparticles in nanofilms of polyelectrolytes will be discussed.

OVERVIEW

Semiconducting Nanoparticles

Nanoparticulate systems are deemed technologically important in the design of miniaturized, ultrahigh density integrated circuits and information storage devices of the future and in some sense an alternate route to overcoming the 100-nm barrier in conventional electronics and develop nanodimensional molecular electronic devices.[10–13]

Owing to their semiconducting properties, group II–VI and IV–VI compounds have been intensively examined by various authors.[14] The size-dependent optoelectronic properties of nanoparticles (diameter 1–100 nm) are attributed to quantum confinement effects.[1–13] Briefly, electronic excitation in semiconductors arises from an exciton (an electron and hole bound pair) localized in a potential well. Theoretical calculations[6–13] have demonstrated that when particle sizes corresponding to the De Broglie wavelength of the free charge carriers are approached, quantum confinement effects become dominant. One manifestation of such an effect is an increase in the optical band gap energy with decreasing particle size that is readily manifested as sharp changes in color visible to the naked eye. Depending on the particle size, the color of CdS colloids may vary from blue to red, whereas PbS nanoparticles may appear pale yellow, orange, red, or black.

Nanoparticles has been studied in micelles,[15–17] vesicles,[18,19] sol–gel glasses,[20,21] zeolites,[22,23] Langmuir–Blodgett (LB) films,[24,25] and polymers.[26,27] In most cases the clusters have poorly defined surfaces and a broad distribution of particle sizes. Empedocles et al.[28] have demonstrated control in preparing monodisperse CdSe clusters using a synthesis medium consisting of trioctylphosphine and its oxide. While size exclusion chromatography permits a narrow size distribution of particle sizes, only minute quantities of the materials are obtained, which is unsuitable for any large-scale applications. On the contrary, the synthetic route used by Empedocles et al.[28] permits the production of gram quantities of nanoparticulate solids. Chemical reactions initiated within the microscopic cavities of zeolites, glasses, polymers, and micelles provide another process of preparing nanoparticles. The shape and sizes of the nanoparticles in these "microreactors" are largely controlled by the restricted geometry of the cavities in which nucleation and growth of these particles occur. Thermodynamic and entropic requirements also play a crucial role in determining the size of these clusters.

Iron Oxide Nanoparticles

In the attempt to improve magnetic recording technology, great effort has been made for obtaining high-density recording media. For this purpose, the production of nanoparticles that are uniform, highly dispersible, and oriented in a matrix is essential. Controlled coercivity (between 500 and 1500 Oe) and high saturation magnetization are important tasks. Increasing the coercivity to an applicable level has been fulfilled by doping or coating the maghemite particles with Co.[29] The production of nanoparticles has been achieved on the basis of a mimetic approach, i.e., with the utilization of an organic support that plays an essential role in crystal nucleation and in its growth control.

Nguyen and Diaz reported a simple synthesis of bulk poly(pyrrole-*N*-propylsulfonate) polymer composites

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120024799

containing nanosized magnetite (γ-Fe_2O_3) particles.[30] Using $FeCl_3$, they polymerized pyrrole-*N*-propylsulfonate, resulting in a black polymeric powder. Using sulfonate groups as nucleation sites for the growth of nanoparticles was a new technique introduced by Ziolo et al.[31,32] when they reported a matrix-mediated synthesis of maghemite (γ-Fe_2O_3). The bulk magnetic material that they synthesized, a γ-Fe_2O_3/polymer nanocomposite, was optically active at room temperature. Using $FeCl_3$ and $FeCl_2$ they ion-exchanged sulfonated polystyrene cross-linked with divinylbenzene which yielded a polymer with iron cations ionically bonded to sulfonate groups. Washing with NaOH, then heating to 60°C, while adding aqueous H_2O_2 solution, oxidized the iron to γ-Fe_2O_3. Ziolo et al. proposed that the sulfonate groups of the polymer matrix provide spatially located sites for the growth of γ-Fe_2O_3, while the void volume in the cross-linked resin imposes a limit on the maximum size of the crystals grown, thereby minimizing aggregation of the iron oxide particles. Particles ranged in size from 5 to 10 nm, as determined by transmission electron microscopy (TEM).

A similar technique was used to prepare a superparamagnetic form of goethite in pores of sulfonated, highly cross-linked poly(divinylbenzene) microspheres.[33] The synthesized iron oxide had a calculated magnetic susceptibility, which was about 3 orders of magnitude larger than bulk goethite. Chemically, ferrous ion oxidation reaction leads to the formation of γ-Fe_2O_3/polymer composites, using a commercial ion exchange resin consisting of sulfonated, lightly cross-linked polystyrene. The use of polymers with different pore diameters and grade of cross-linkage lead to the oxidation of ferrous chloride to different iron oxides. In this behavior it could be shown that organic matrices have a dramatic influence on the crystallization product.

Biomineralization

Biomineralization occurs naturally in many biological systems including bacteria.[4] In the presence of a supersaturated solution, minerals may form on the surface of a living organism. The formed structures are characterized by a high degree of regularity and the supramolecular architectures are finely controlled. Although biomineralization had been widely studied,[4,34–36] the mechanisms of growth are not always understood. The supramolecular organization of the organic support seems to have an essential role in the process; these organized architectures (such as vesicles, micelles, polymeric networks) provide functionalized surfaces that act as templates and assist in the interfacial molecular recognition. In many cases the organic matrix provides precise sites for oriented nucleation.

LAYER-BY-LAYER DEPOSITION

The layer-by-layer deposition technique of building supramolecular multilayers on solid substrates by adsorbing polyelectrolytes has emerged as a simple means of

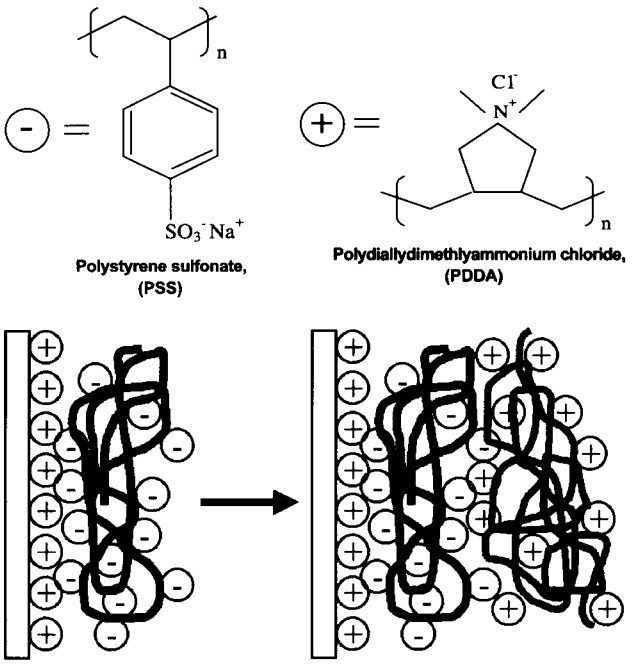


Fig. 1 Schematic representation of layer-by-layer deposition for a positively charged surface.

producing templates of controlled thickness. Decher[37,38] has established that such multilayers may be reproducibly built and the resultant structure is mechanically and thermally stable. Such films have been utilized in the preparation of electroluminescent devices, zener diodes, and the possibility of their use in flat screen displays has been suggested.[13] The porous and supramolecular structure of these films provides opportunities for studying chemical reactions in the nanoscale regime. Fig. 1 shows pictorially an example of the layer-by-layer deposition technique for sequential deposition of a polyanion followed by a polycation on a positively charged surface.

The procedure of layer-by-layer deposition is simple. For example, poly(diallyldimethylammoniumchloride) (PDDA), a polycation, and poly(styrenesulfonate) (PSS), a polyanion, can be deposited sequentially on a negatively charged surface as follows. The polyionic solutions of PDDA and PSS can be prepared separately in 20-mM concentrations: the former is dissolved in deionized water, whereas the latter is dissolved in a 0.1-M NaOH solution, and then adjusted to a pH of 4.5 by drop wise addition of 0.1 M HCl. All concentrations are calculated based on the monomer molecular weights. Polymer films can be assembled on negatively charged substrates if one starts with PDDA. A PDDA/PSS deposition cycle is termed the deposition of one layer-pair. For nanoparticle nucleation, deposition is preferably ended with the first half of a layer-pair in order to terminate the film with a PDDA layer, as will be discussed later. The deposition of the layer-by-layer films can be followed by UV–visible spectroscopy as shown in Fig. 2, and the absorbance is linear with the number of layer-pairs deposited. For the above conditions each layer-pair is approximately 4–5 nm thick, so that a 10.5-layer-pair-thick nanofilm is about 50 nm in total thickness.

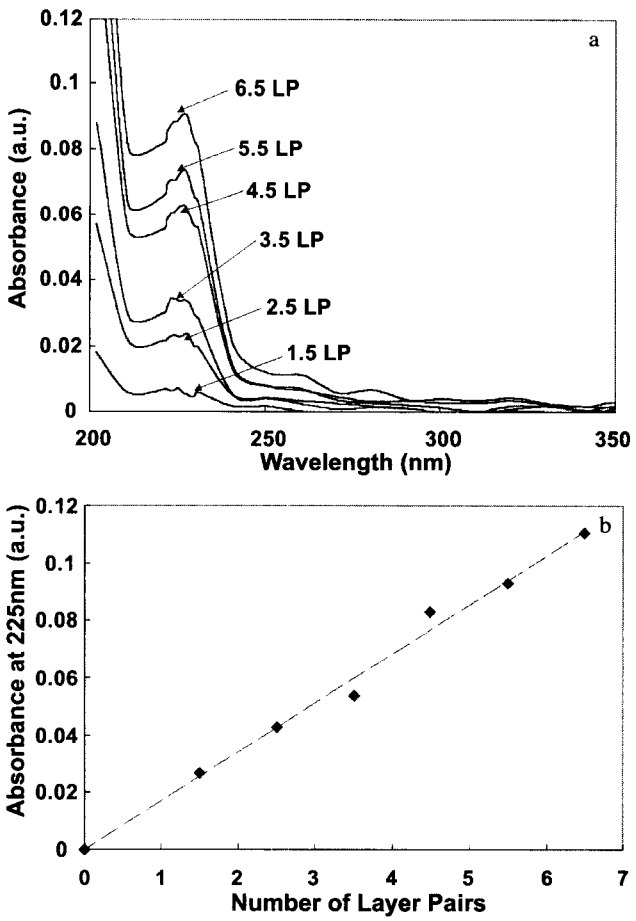


Fig. 2 a) UV–visible absorption for a number of layer-pair (LP) deposition (without absorption–oxidation cycles); b) absorbance at 225 nm with respect to the number of layer-pairs deposited. (From Fojas et al. [Ref. 39], with permission.)

NUCLEATION OF NANOPARTICLES IN NANOFILMS OF POLYELECTROLYTES

It has been demonstrated by Stroeve and coworkers[39–43] that divalent and trivalent ions can bind to the negatively charged sulfonate groups in layer-by-layer polyelectrolyte nanofilms and that the absorbed ions can be oxidized to form nanoparticles of oxides, or reduced to form nanoparticles or other compounds. For example, nucleation of metal oxide nanoparticles within a nanofilm of polyelectrolyte occurs by cycling the polymer-coated substrates first in M^{2+} and then in NaOH or NH_4OH solutions (absorption–oxidation cycles or oxidative hydrolysis cycles). The M^{2+} solutions can be prepared in concentrations of 4–40 mM by dissolving a M^{2+} salt [e.g., MCl_2 or $M(NO_3)_2$] in purified water. The solutions of 0.1–1.0 M base are made using purified water. Before cycling, all solutions are vigorously sparged for 45 min with nitrogen gas to remove any dissolved oxygen in solution. In Schlenken tubes, under nitrogen gas, the polymer nanofilm on a substrate can be exposed to the M^{2+} solution for a few minutes. Typical substrates can be quartz, silicon wafers, ZnSe wafers, membranes, or more complex-shaped substrates such as catalytic particles. The substrates are rinsed with degassed and purified water to remove excess ions absorbed in the film, and then exposed to NaOH or NH_4OH solution for several minutes. Substrates are then rinsed with degassed and purified water and dried completely before the absorption–oxidation cycle is repeated. Fig. 3 shows that the process of oxidative nucleation and growth can be followed by UV–visible spectroscopy.[39] Fig. 3b shows that the nucleation of nickel hydroxide nanoparticles with the number of oxidative hydrolysis cycles is linear. The resulting nickel hydroxide nanoparticles after x cycles is shown in Fig. 4. Each oxidative hydrolysis cycle nucleates new nanoparticles, while nanoparticles from previous cycles grow in size.

Dante et al.[40] demonstrated the formation of akaganéite (β-FeOOH) in multilayered, polymer films of PDDA and PSS produced by the layer-by-layer technique. Ferrous chloride was used as the starting metal ion solution. Uniform, needlelike nanoparticles of β-FeOOH were obtained inside the nanofilm. Transmission electron microscopy showed uniform needles with a diameter of 10 nm and a length of 100 nm. Deposition of the nanofilms had to be terminated with the last layer being PDDA. In the case that the last layer was PSS, two types of nanoparticles were obtained: β-FeOOH and γ-Fe_2O_3. The γ-Fe_2O_3 nanoparticles were spherical and 10 nm in size. Presumably, the mechanism of oxidation of iron ions on the outermost PSS surface layer was different from the oxidation of the iron ions bound to the inner PSS layers.

Dutta et al.[41] studied the nucleation and growth of lepidocrocite (γ-FeOOH) crystals in a nanofilm generated as a result of a layer-by-layer adsorption of PDDA and PSS. Interest in lepidocrocite stems from the fact that it is paramagnetic at room temperature having a low Neel temperature of 77 K and is easily converted into other industrially important magnetic oxides, namely, maghemite and hematite upon heating. Nanothin films were formed by the layer-by-layer deposition technique on a quartz substrate. Nucleation of nanoparticles in the nanofilm of polyion was initiated by adsorbing ferric nitrate and its subsequent hydrolysis with ammonium hydroxide. Repeating the above process resulted in an increase in the density of the nanoparticles initially formed followed by the appearance of crystallites that grow in dimensions with the number of adsorption and hydrolysis cycles. The size of the crystals can be controlled by the number of cycles used. Analysis of the steady-state UV–visible absorption spectra of the films revealed the formation of lepidocrocites, which was confirmed by FTIR and selected area electron diffraction (SAED) studies. An important feature of this work was that lepidocrocite (γ-FeOOH) is formed instead of akaganéite, which was generated when ferrous chloride was used as the starting material as reported in the work of Dante et al.[40] It therefore appears likely that the initial starting material plays a key role in determining the structural and morphological characteristics of iron oxyhydroxides although their chemical compositions are the same.

Nucleation and growth of lead sulfide (PbS) nanoparticles in a PDDA–PSS film produced by the layer-by-layer deposition technique has been studied by Dutta et al.[42] Interest in PbS arises from the fact that it is a

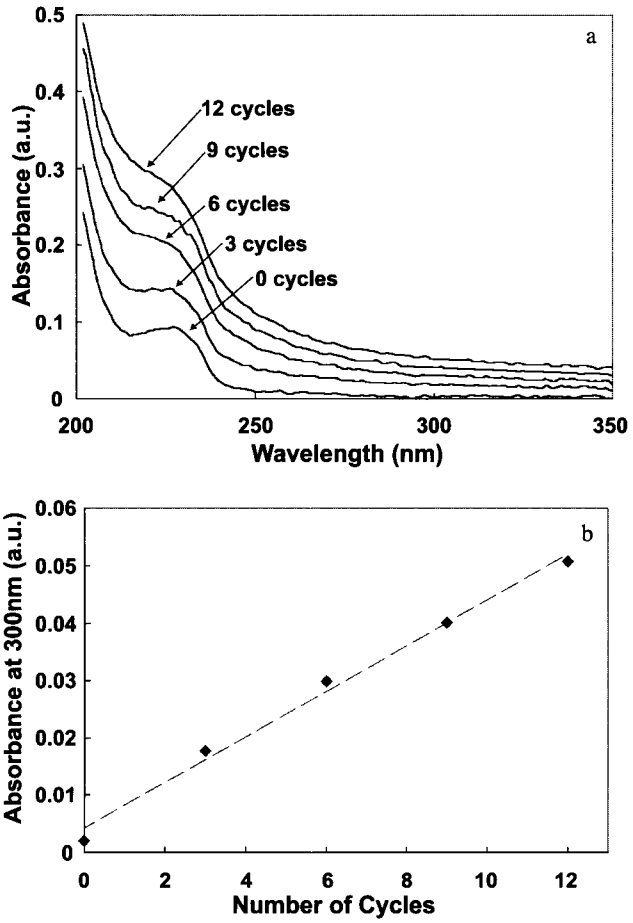


Fig. 3 a) UV–visible absorption spectra for 6.5 layer-pairs cycled in a 40-mM $NiCl_2 \cdot 6H_2O$ solution and in 1.0 M NaOH solution; b) absorption at 300 nm with respect to number of oxidative cycles. (From Fojas et al. [Ref. 39], with permission.)

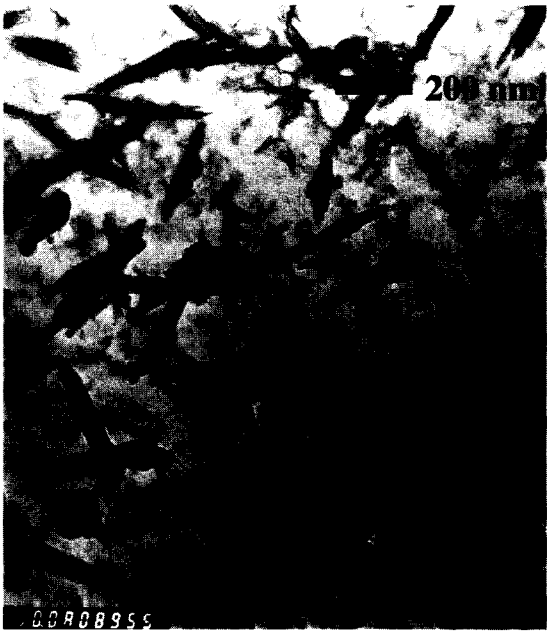
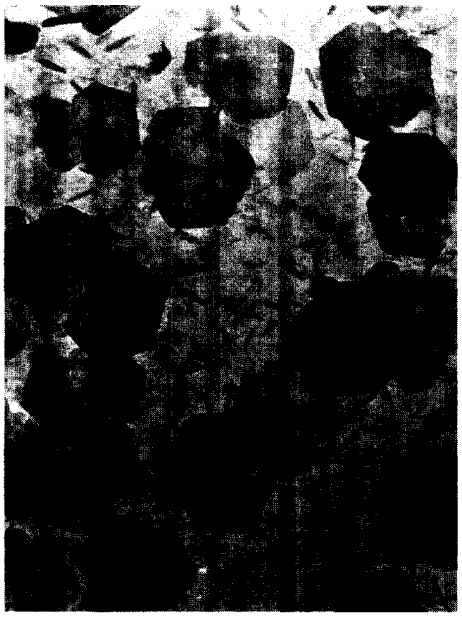


Fig. 4 TEM image of 2.5 layer-pairs with eight oxidative hydrolysis cycles using 4 mM $NiCl_2 \cdot 6H_2O$ and ammonia base (50 K × magnification). The smaller particles are about 50 nm in length, while the larger particles are about 300 nm in length. The smaller nanoparticles are generated with each cycle, whereas existing particles continue to grow in size with each cycle. (From Fojas et al. [Ref. 39], with permission.)

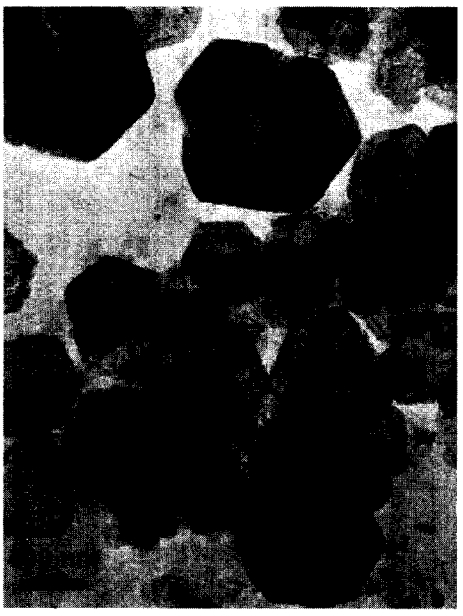
(a)



(b)

Fig. 5 TEM micrographs of a 3.5 layer pairs of PDDA–PSS films at different stages of the absorption–hydrolysis process in a nitrogen-enriched environment: (a) two cycles, and in an oxygen-enriched microenvironment panels: (b) four cycles (c) eight cycles. The concentrations of cobalt chloride and sodium hydroxide solutions were 4 and 10 mM, respectively. (From Zhang et al. [Ref. 44], with permission.)

semiconductor having a small band gap (0.41 eV) and large exciton diameter (18 nm) that permits size-quantization effects to be observable even for large-sized particles or crystallites. Moreover, the nonlinear optical (NLO) properties of PbS nanoparticles show large differences in their optical limiting behavior below and above the absorption edge, suggesting that such systems may be utilized in high-speed switching.[43] Chemical reaction within the polymer film was initiated by Dutta et al.[42] by absorbing Pb^{2+} from an aqueous solution of $Pb(NO_3)_2$ followed by exposing the film to H_2S gas. Electron microscopic examination of the films revealed that while nanoparticles are formed in films that were subject to one or two reaction cycles, large crystallites were formed when these films were exposed to a large number[10] of reaction cycles. In the latter case a broad distribution of particle sizes is observed and may perhaps be attributed to a form of Ostwald ripening. Detailed studies show the nucleation and growth of the PbS nanoparticles into larger crystallites. UV–visible absorption studies reveal that the absorption spectral profiles of the films are dependent on the size of the PbS crystallites. The broadened absorption spectral profile observed for films subject to a large number of reaction cycles may be attributed to the superposition of the spectral profiles of the small clusters that tend to be blue shifted due to quantum confinement effects and the large clusters that are red shifted. Crystal size can be controlled by the number of cycles. Essentially, a single cycle only gives nanoparticles that continue to grow in size with more cycles. Instead of using a reducing reaction, it was also possible to oxidize the Pb ions to make $PbSO_4$ particles.[42]



(c)

Fig. 5 (*Continued.*)

Zhang et al.[44] reported the oxidative hydrolysis of Co^{2+} ions absorbed in organized, multilayered, polymer films to form cobalt hydroxide nanocrystals. It was found in this study that using more polymer layer-pairs result in more crystal growth. Hydrolysis of the Co^{2+} ions in a

nitrogen-rich environment gave rise to mainly needlelike crystallites of α-$Co(OH)_2$ that were initially about 100 nm in length and then increased in size with the number of absorption–hydrolysis cycles. However, in an oxygen-rich microenvironment, hexagonal crystallites were found to be predominant. X-ray diffraction (SAED) and TEM studies revealed that these hexagonal crystallites are mainly that of β-$Co(OH)_2$. Fig. 5 shows the transition from the alpha to the beta form as observed by TEM. These studies suggest that by varying the nitrogen-to-oxygen ratio the formation of α-$Co(OH)_2$ and β-$Co(OH)_2$ may be controlled.

Rubner and colleagues[45,46] have further improved the method of Dante et al.[40] by using weak polyelectrolytes, such as poly(acrylic acid) (PAA), instead of a strong polyelectrolyte such as PSS. The advantage of their technique is that the linear charge density can be varied with pH. Thus it is possible to control the charge density of the adsorbing polyelectrolyte and the charge of the previously adsorbed polymer. This technique allows the control of the thickness of the nanofilm, the composition, surface properties, and the level of polymer interpenetration. Rubner and colleagues synthesized layer-by-layer nanofilms containing silver nanoparticles and semiconductor nanoparticles. They showed that it was possible to stratify the nanofilm in layers with nanoparticles and layers without nanoparticles. Dai and Bruening used postreduction of metal ions to obtain layer-by-layer films of catalytic metal nanoparticles with a size range of 4 to 30 nm.[47] Recently, Rubner and colleagues nucleated palladium nanoparticles as catalytic seeds for further growth of electroless nickel on the nanoparticle seeds.[48] The size of the seeded nickel nanoparticles could be controlled and up to 14-nm-diameter nanoparticles were obtained inside the nanofilms.

CONCLUSION

The layer-by-layer technique of depositing polyions on substrates to build nanofilms of specific thickness, composition, and surface charge can be used to carry out nanoreaction inside the films to form nanoparticle–polymer complexes. Nanofilms of nanoparticle–polymer are useful in coatings, catalysis, electrical, optical, magnetic, electrooptical, and magnetooptical applications. Manipulation of film type and reaction conditions allows for control of the nanoparticle product, distribution, and size inside the polymer matrix.

REFERENCES

1. Nirmal, M.; Brus, L. Luminescence photophysics in semiconductor nanocrystals. Acc. Chem. Res. **1999**, *32*, 407.
2. Alivisatos, A.P. Semiconductor clusters, nanocrystals, and quantum dots. Science **1996**, *271*, 933.
3. Chan, W.C.W.; Nie, S. Quantum dot bioconjugates for ultrasensitive nonisotopic detection. Science **1998**, *281*, 2016.
4. Dujardin, E.; Mann, S. Bio-inspired materials chemistry. Adv. Mater. **2002**, *11*, 775–788.
5. Brus, L.E. Luminescence of direct and indirect gap quantum semiconductor crystallites. MRS Symp. Ser. **1992**, *272*, 215.
6. Wang, Y. Nonlinear optical properties of nanometer-sized semiconductor clusters. Acc. Chem. Res. **1991**, *24*, 133.
7. Steigerwald, M.L.; Brus, L.E. Semiconductor crystallites: A class of large molecules. Acc. Chem. Res. **1990**, *23*, 183.
8. Norris, D.J.; Bawendi, M.G.; Brus, L.E. Optical Properties of Semiconductor Nanocrystals (Quantum Dots). In *Molecular Electronics*; Jortner, J., Ratner, M., Eds.; Blackwell: Oxford, U.K., 1997.
9. Weller, H. Quantisized semiconductor particles—A novel state of matter for materials science. Adv. Mater. **1993**, *5*, 88.
10. Koyama, H.; Araki, M.; Yamamoto, Y.; Koshida, N. Visible photoluminescence of porous silicon and related optical properties. Jpn. J. Appl. Phys. **1991**, *30*, 3606.
11. Dabbousi, B.O.; Bawendi, M.G.; Onitsuka, O.; Rubner, M.F. Electroluminescence from CdSe quantum-dot/polymer composites. Appl. Phys. Lett. **1995**, *66*, 1316.
12. Nirmal, M.; Dabbousi, B.O.; Bawendi, M.G.; Macklin, J.J.; Trautman, J.K.; Harris, T.D.; Brus, L.E. Fluorescence intermittency in single cadmium selenide nanocrystals. Nature **1996**, *383*, 802.
13. Cassagneau, T.; Mallouk, T.E.; Fendler, J.H. Layer-by-layer assembly of thin film zener diodes from conducting polymers and CdSe nanoparticles. J. Am. Chem. Soc. **1998**, *120*, 7848.
14. *Semiconductor Nanoclusters—Physical, Chemical and Catalytic Aspects*; Kamat, P.V., Meisel, D., Eds.; Elsevier: Amsterdam, 1997.
15. Pileni, M.P. *Structure and Reactivity in Reverse Micelles*; Elsevier: Amsterdam, 1989.
16. Petit, C.; Lixon, P.; Pileni, M.P. Synthesis of cadmium sulfide in situ in reverse micelles: 2. Influence of the interface on the growth of the particles. J. Phys. Chem. **1990**, *94*, 1598.
17. Fendler, J.H. Atomic and molecular clusters in membrane mimetic chemistry. Chem. Rev. **1987**, *87*, 877.
18. Youn, H.C.; Baral, S.; Fendler, J.H. Dihexadecyl phosphate, vesicle-stabilized and in situ generated mixed cadmium sulfide and zinc sulfide semiconductor particles: Preparation and utilization for photosensitized charge separation and hydrogen generation. J. Phys. Chem. **1998**, *92*, 6320.

19. Kennedy, M.T.; Korgel, B.A.; Montbouquette, H.G.; Zasadinski, J.A. Cryo-transmission electron microscopy confirms controlled synthesis of cadmium sulfide nanocrystals within lecithin vesicles. Chem. Mater. **1998**, *10*, 2116.
20. Rajh, T.; Micic, O.I.; Lawless, D.; Serpone, N. Semiconductor photophysics: 7. Photoluminescence and picosecond charge carrier dynamics in cadmium sulfide quantum dots confined in a silicate glass. J. Phys. Chem. **1992**, *96*, 4633.
21. Minti, H.; Eyal, M.; Reisfeld, R.M.; Berkovic, G. Chem .Phys. Lett. **1991**, *183*, 277.
22. Wang, Y.; Herron, N. Photoluminescence and relaxation dynamics of cadmium sulfide superclusters in zeolites. J. Phys. Chem. **1988**, *92*, 4988.
23. Ozin, G.A.; Steele, M.R.; Holmes, A.J. Intrazeolite topotactic MOCVD. 3-Dimensional structure-controlled synthesis of II–VI semiconductor nanoclusters. Chem. Mater. **1994**, *6*, 999.
24. Yang, J.; Meldrum, F.C.; Fendler, J.H. Epitaxial growth of size-quantized cadmium sulfide crystals under arachidic acid monolayers. J. Phys. Chem. **1995**, *99*, 5500.
25. Yang, J.; Fendler, J.H. Morphology control of PbS nanocrystallites, epitaxially grown under mixed monolayers. J. Phys. Chem. **1995**, *99*, 5505.
26. Tassoni, R.; Schrock, R.R. Synthesis of PbS nanoclusters within microphase-separated diblock copolymer films. Chem. Mater. **1994**, *6*, 744.
27. Kane, R.S.; Cohen, R.E.; Silby, R.J. Synthesis of PbS nanoclusters within block copolymer nanoreactors. Chem. Mater. **1996**, *8*, 1919.
28. Empedocles, S.A.; Norris, D.J.; Bawendi, M.G. Photoluminescence spectroscopy of single CdSe nanocrystallite quantum dots. Phys. Rev. Lett. **1996**, *77*, 3873.
29. Saito, S. *Fine Ceramics.* Elsevier: New York, 1988.
30. Nguyen, M.T.; Diaz, A.F. A novel method for the preparation of magnetic nanoparticles in a polypyrrole powder. Adv. Mater. **1994**, *6* (11), 858–860.
31. Ziolo, R.F.; Giannelis, E.P.; Weinstein, B.A.; O'Horo, M.P.; Ganguly, B.N.; Mehrotra, V.; Russell, M.W.; Huffman, D.R. Matrix-mediated synthesis of nanocrystalline γ-ferric oxide: A new optically transparent magnetic material. Science **1992**, *257*, 219–223.
32. Vassiliou, J.K.; Mehrotra, V.; Russell, M.W.; Giannelis, E.P.; McMichael, R.D.; Shull, R.D.; Ziolo, R.F. Magnetic and optical properties of γ-Fe_2O_3 nanocrystals. J. Appl. Phys. **1993**, *73* (10), 5109–5115.
33. Winnik, F.M.; Morneau, A.; Ziolo, R.F.; Stoever, H.D.H.; Li, W.H. Template-controlled synthesis of superparamagnetic goethite within macroporous polymeric microspheres. Langmuir **1995**, *11*, 3660–3666.
34. Tarasevich, B.J.; Rieke, P.C.; Liu, J. Nucleation and growth of oriented ceramic films onto organic interfaces. Chem. Mater. **1996**, *8*, 292–300.
35. Lowenstam, H.A.; Weiner, S. *On Biomineralization*; Oxford University Press: Oxford, 1989.
36. Mann, S.; Webb, J.; Williams, R.J.P. *Biomineralization: Chemical and Biochemical Perspectives*; VCH: Weinheim, 1989.
37. Decher, G. *Comprehensive Supramolecular Chemistry*; Sauvage, J.P., Ed.; Pergamon Press: New York, 1996.
38. Decher, G. Fuzzy nanoassemblies: Toward layered polymeric multicomposites. Science **1997**, *277*, 1232.
39. Fojas, A.M.; Murphy, E.; Stroeve, P. Layer-by-layer polymeric supramolecular structures containing nickel hydroxide nanoparticles and microcrystallites. Ind. Eng. Chem. Res. **2002**, *41*, 2662–2667.
40. Dante, S.; Hou, Z.; Risbud, S.; Stroeve, P. Nucleation of iron oxy-hydroxide nanoparticles by layer-by-layer polyionic assemblies. Langmuir **1999**, *15*, 2176.
41. Dutta, A.; Jarero, G.; Zhang, L.; Stroeve, P. In-situ nucleation and growth of γ-FeOOH nanocrystallites in polymeric supramolecular assemblies. Chem. Mater. **2000**, *12*, 176.
42. Dutta, A.; Ho, T.; Zhang, L.; Stroeve, P. Nucleation and growth of lead sulfide nano- and microcrystallites in supramolecular polymer assemblies. Chem. Mater. **2000**, *12*, 1042.
43. Colvin, V.L.; Schlamp, M.C.; Alivisatos, A.P. Light-emitting diodes made from cadmium selenide nanocrystals and a semiconducting polymer. Nature **1994**, *370*, 354.
44. Zhang, L.; Dutta, A.; Jarero, G.; Stroeve, P. Nucleation and growth of cobalt hydroxide crystallites in organized polymeric multilayers. Langmuir **2000**, *16*, 7095.
45. Joly, S.; Kane, R.; Radzilowski, T.; Wang, A.; Cohen, R.E.; Thomas, E.L.; Rubner, M.F. Multilayer nanoreactors for metallic and semiconducting particles. Langmuir **2000**, *16*, 1354–1359.
46. Wang, T.C.; Rubner, M.F.; Cohen, R.E. Polyelectrolyte multilayer nanoreactors for preparing silver nanoparticle composites: Controlling metal concentration and nanoparticle size. Langmuir **2002**, *18*, 3370–3375.
47. Dai, J.; Bruening, M.L. Catalytic nanoparticles formed by reduction of metal ions in multilayered polyelectrolyte films. Nano Lett. **2002**, *2*, 497–500.
48. Wang, T.C.; Rubner, M.F.; Cohen, R.E. Manipulating nanoparticle size within polyelectrolyte multilayers via electroless nickel deposition. Chem. Mater. **2003**, *15*, 299–304.

Nucleoside- and Nucleobase-Substituted Oligopyrrolic Macrocycles

Vladimír Král
Martin Valík
Tatiana V. Shishkanova
Institute of Chemical Technology, Prague, Czech Republic

Jonathan L. Sessler
The University of Texas at Austin, Austin, Texas, U.S.A.

INTRODUCTION

This chapter summarizes the design and synthesis of conjugates resulting from the combination of nucleic acid base derivatives, particularly nucleosides and nucleotides, with oligopyrrole macrocycles. In addition to detailing the methods of preparation and fundamental chemical properties, the vision behind such projects and the various applications targeted for use by these conjugates are also described. These systems combine into one chemical entity two of the more important biological building blocks found in Nature, namely, nucleic acid bases (nucleobases) and porphyrins, and this leads to new constructs with rather unique chemical features as well as novel molecular recognition properties. Although the focus of the present chapter will be on systems containing natural nucleobase components, it is important to appreciate that the porphyrin portion of these constructs can include a range of synthetic oligopyrrole macrocycles that have no biological equivalent. This gives rise to products with unexpected supramolecular properties, including those associated with anion recognition.

BACKGROUND

Nucleotides and nucleic acids play key roles in many biological processes ranging from information processing to energy storage and transduction. In addition, a number of nucleotide and nucleoside analogues exhibit antiviral activity. As a consequence, considerable study is devoted to simple, monomeric nucleic acid bases (''nucleobases''; NBs), both in terms of understanding their fundamental properties and in terms of developing potential new drug leads. Activity is also high in the areas of synthesis (e.g., new methods development) and sensing (i.e., finding ways to detect specific nucleic acid derivatives selectively). The NBs also inspired supramolecular chemists as a result of their ability to undergo complementary Watson–Crick-type base pairing. This, of course, represents the most ubiquitous and fundamental kind of hydrogen bond-based molecular recognition. It thus provided an incentive to construct new, rationally designed conjugates, including covalent combinations of nucleobase with oligopyrrolic macrocycles (OPMC). These latter systems, which are the focus of this review, possess unique photophysical and molecular recognition properties that are specifically defined by the nature of the constituent NB and OPMC components, as well as the extent of ''chemical communication'' between two parts of what may be considered a complex, multifunctional receptor molecule. In this entry the chemistry and properties of NB–OPMC conjugates will be reviewed with a focus on the specific choice of NB, OPMC, and/or covalent linkage. However, to keep the discussion within reasonable limits, only naturally occurring purine and pyrimidine NB component(s) will be considered. By contrast, our coverage of known systems containing various linking spacer (SP) subunits and OPMCs attempts to be more comprehensive. Of the latter, which run the gamut from biomimetic to wholly synthetic, porphyrins have received the greatest attention. Thus we start our discussion with NB–SP–OPMC conjugate systems that are based on this biologically all-important chromophore.

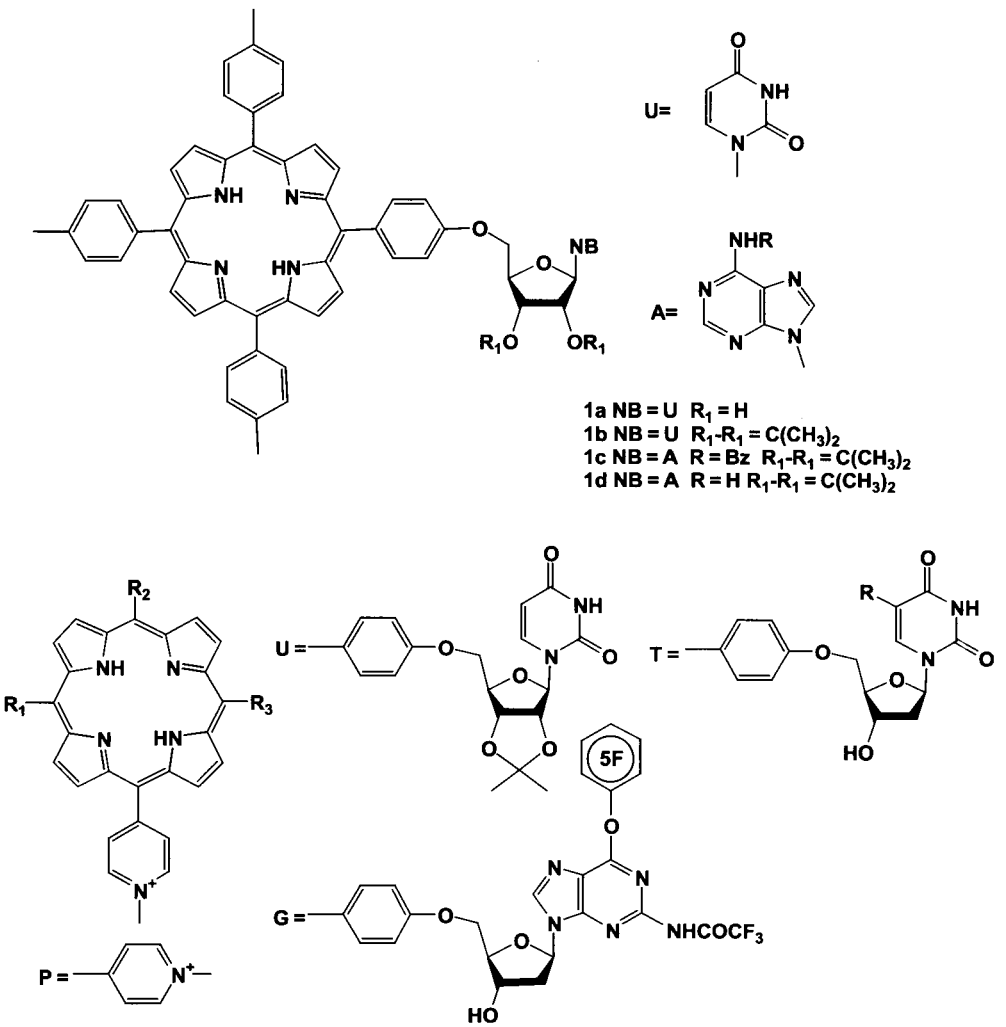
Nucleobase–Porphyrin Derivatives

The first porphyrin nucleoside conjugates, represented by structures **1**, were synthesized by Kus and coworkers in 1990.[1] Unfortunately, these systems proved essentially insoluble in water,[2] a feature that diminished their biomedical significance and limited their utility in terms of various targeted therapeutic applications. To overcome this problem, the *meso*-(1-methylpyridinium-4-yl)$_n$porphyrin core was chosen as the key OPMC subunit. This choice was dictated by solubility considerations and by the recognition that *N*-methylated pyridyl porphyrins and related systems bind strongly to DNA with a preference of

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120018316

G–C base pair regions[3] and that certain methylated derivatives can be photoactivated to generate active species, such as single oxygen, that are known to cleave oligonucleotides.[4] The remaining *meso* positions were then used to attach the nucleoside parts. In this way, the first water soluble porphyrinyl nucleosides **2** were prepared, specifically via the condensation of *meso*-(pyridin-4-yl)$_n$-(*p*-hydroxylphenyl)$_{4-n}$porphyrins (n=1–3) with 5′-*O*-tosylate of 2′,3′-*O*-isopropylideneuridine and subsequently subjecting the pyridine cores to methylation.[2,5] The same synthetic strategy was used to prepare analogues containing other porphyrinyl–nucleosides, in particular the thymidine derivatives, **3**,[6] the fluorinated thymidine (**4a**, **4b**),[7] and guanosine (**4c**) porphyrinyl derivatives.[7] The biological properties of the porphyrinyl–nucleoside conjugates **2a** and **3** and their porphyrin-centered Co(II) complexes were then studied in terms of growth suppression of malignant melanoma cells.[6] Significant suppression was achieved using the Co(II) complex of porphyrinyl–dithymidine **3b**. In this instance, the extent of inhibition reached 95%. By contrast, in the case of the corresponding Co(II) complex of **3a**, containing one thymidine unit, only a low level of suppression was observed[6] (Scheme 1).

A strategy analogous to that described above was also used to prepare the two adenine-containing porphyrinyl–



Scheme 1

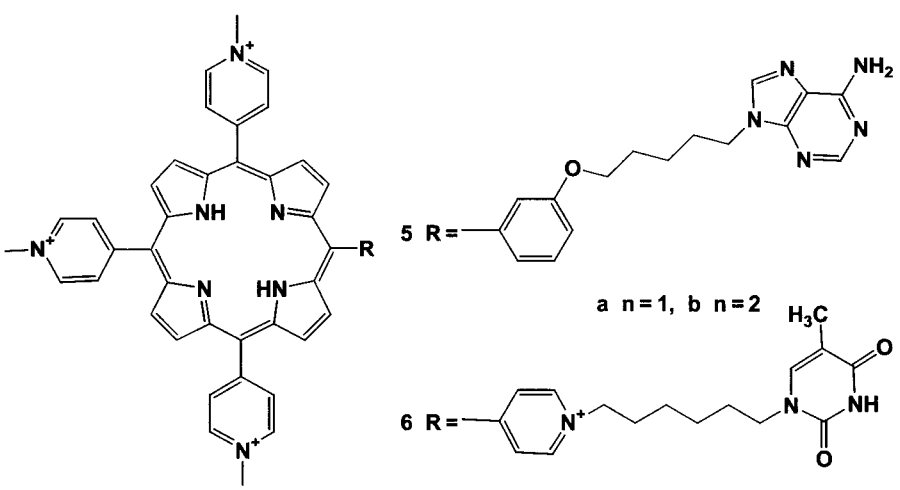
NB derivatives[8] **5a** and **5b**. The synthesis was modified, however, such that:

i) The *N*-methylation step was carried out prior to forming the hybrid molecule.
ii) The bromo derivative of the nucleobase-functionalized alkyl spacer was used instead of the tosylate employed in the case of more tightly coupled conjugates **1**–**4**. Unfortunately, this procedure failed to provide a more loosely linked thymidine analogue of **3a** (e.g., **6**). Thus an alternative approach was developed. It involved the treatment of 5,10,15-tris(1-methylpyridinium-4-yl)-20-(pyridin-4-yl)porphyrin with a 6-carbon bromoalkyl derivative of thymidine; this gave target **6** in almost quantitative yield.[8]

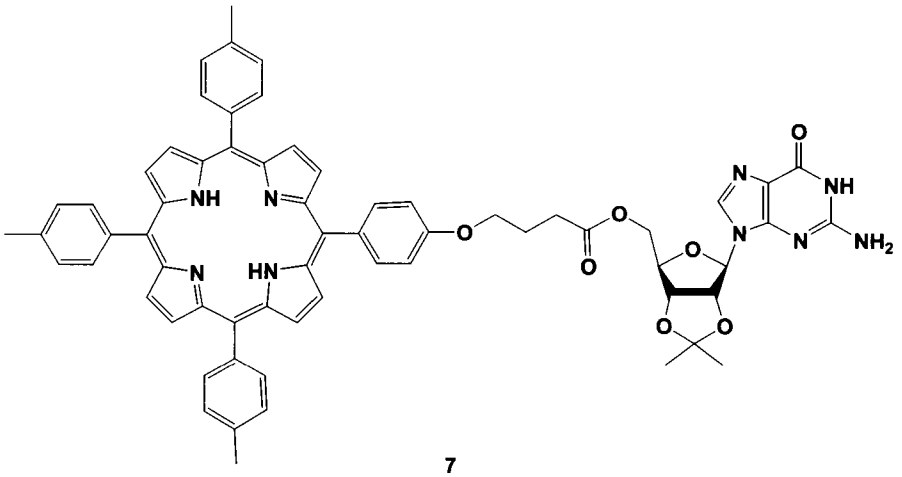
Once in hand, the interactions of **5** and **6** with nucleosides and polynucleotides were studied via UV–VIS and fluorescence-based spectroscopic titration methods. The calculated K_S values (1:1) for all combination of **5** and **6** with simple nucleosides were the same within the error of the method (ca 1.6×10^4 M^{-1}). In other words, no preference was observed between the porphyrin–nucleobase conjugates and the corresponding complementary Watson–Crick nucleotides. The absence of any significant Watson–Crick-type hydrogen bonding selectivity was interpreted as reflecting the dominance of intermolecular stacking interactions, as opposed to base-pairing ones. In the case of polynucleotide, two binding processes were inferred from the fluorimetric titration experiments. First, a set of spectroscopic changes at $r>0.1$ (r=ratio porphyrin/polynucleotide) were attributed to an intermolecular association of the porphyrin portion of the conjugate with the polynucleotide polyanion. Second, a set of different emission features were observed when the intercalation binding sites were present in large excess ($r<0.1$) that were explained in terms of the porphyrin subunits intercalating into the polynucleotide backbone. Interestingly, some selectivities[8] were observed in the case of titrations carried out with polyU, with different results being seen for conjugates **5** containing complementary bases than for the corresponding noncomplementary conjugate **6**[8] (Scheme 2).

With a different focus, Masiero et al. constructed the alkyl-tethered porphyrin-guanosinyl derivative **7**. The propensity of guanosine derivatives to self-assemble into so-called G-quartets was then used to prepare a novel circular porphyrin array.[9] In terms of specifics, these researchers showed that system **7** self-assembles in chloroform solution in the presence of potassium picrate to form a supramolecular complex that consists of eight units of **7** arranged in the form of two G-quartets, as evidenced by UV–VIS, CD, and ^{1}H NMR spectroscopic studies. In the context of this work, complexation of the cation has a dramatic effect on the CD spectrum. In particular, the addition of K^+ leads to the appearance of excitation couplets in both the guanine absorption and porphyrin Soret spectral regions. The couplet observed in the Soret region was deemed particularly significant insomuch as it indicates an intermolecular electronic interaction between the porphyrin chromophores.[9] This, in turn, supports the conclusion that the porphyrins are arrayed unsymmetrically around the central $[G]_8K^+$ "core" (Scheme 3).

Nucleobase–porphyrin constructs containing nucleosides with free ribofuranose hydroxyl groups have also been prepared. For instance, Li and Czuchajowski prepared the porphyrin nucleoside analogues (**8**) that are



Scheme 2

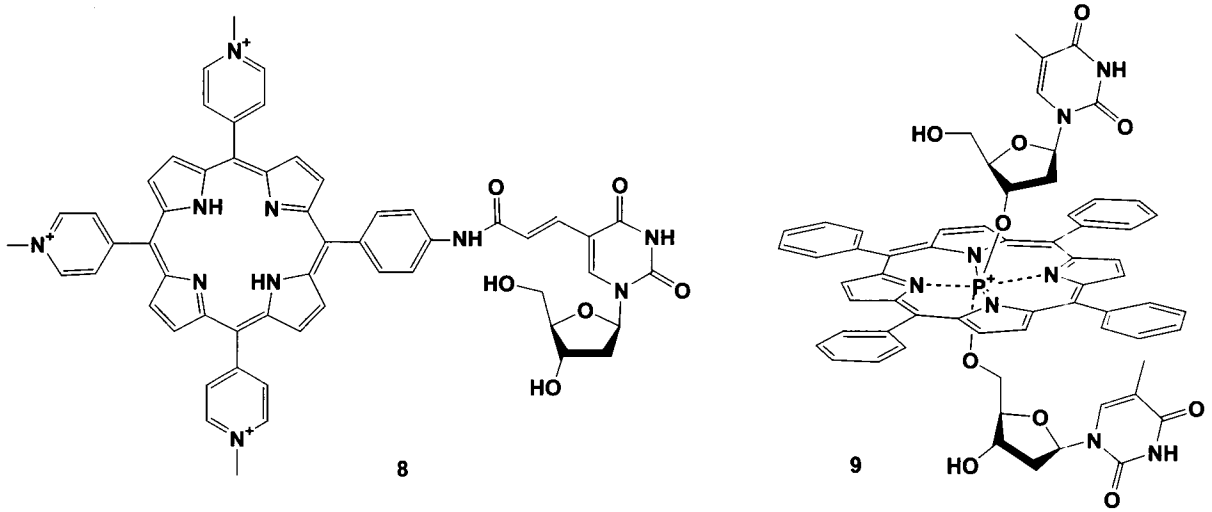


7

Scheme 3

attractive as a potential building block for porphyrin–oligonucletide/DNA systems wherein the number and position of the porphyrin subunits can, at least in principle, be controlled.[10] This system (**8**) was generated from (+)-6-iodo-2′-deoxyuridine and the 5,10,15-tri(4-pyridyl)-20-(*p*-acrylamidophenyl)porphyrin using a Heck reaction. Its interactions with the 20-mer oligonucleotide duplexes $(dG)_{20}$–$(dC)_{20}$ and $(dA)_{20}$–$(dT)_{20}$ were then studied by UV–VIS spectroscopy, as were those with the DNA plasmid-pRIT6. Comparisons with the methylated starting 5,10,15-tri(4-pyridyl)-20-(*p*-acrylamidophenyl)porphyrin revealed that, in every case, the porphyrin Soret band was red-shifted by 6–14 nm. Hypochromicity in the range of 10–54% was also observed. Interestingly, derivative **8** gave rise to a greater degree of bathochromic shift and a lower level of induced hypochromicity than the corresponding 5,10,15-tri(1-methylpyridinium-4-yl)-20-(*p*-acrylamidophenyl)porphyrin "control."[10]

In all of the above examples, the nucleosides were covalently attached to the porphyrin subunit via functionalization of a *meso* aryl substituent. This is not, however, the only way to link these two kinds of components, a point that was elegantly underscored by Goh and Czuchajowski. These researchers bound two nucleosides to the axial positions of phosphorus(V) porphyrins through the ribofuranose 3′-hydroxyl and 5′-hydroxyl substituents[11] The resulting products, potentially a mixture of three *O*-P(V) coordination isomers (vide infra), contain free 3′-hydroxyl and the 5′-hydroxyl groups (cf. structure **9**) and can thus be used, at least in principle, to



8 9

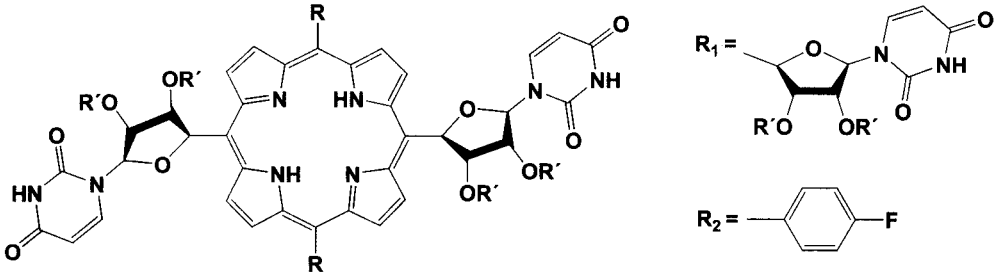
Scheme 4

prepare oligo-porphyrin DNA analogues linked through a positively charged phosphorus(V) porphyrin "backbone." They were prepared by subjecting a dichlorophosphorus(V) tetraphenylporphyrin to nucleophilic substitution using a nonprotected nucleoside as the nucleophile. The least polar and least sterically hindered 5′-,5′-*O*-isomer was formed with the highest yield from the three possible (and observed) products (i.e., the 3′-,3′-, 3′-,5′-, and 5′-,5′-*O*-isomers). However, this isomer proved less stable than the corresponding 3′-,3′- and 3′-,5′-isomers and was found to undergo rapid hydrolysis to the corresponding dihydroxyl phosphorus(V) porphyrin unless protected from light and moisture (Scheme 4).

A different kind of linked system was described by Cornia et al. These workers prepared porphyrin–uridine derivatives wherein the uridine subunits are anchored to the porphyrin macrocycle by means of robust carbon–carbon bonds (i.e., systems **10** and **11**).[12] The synthesis of these systems, which contain direct meso-ribofuranose bonds, began with the coupling of two equiv. of pyrrole to one equiv. of 2′,3′-*O*-isopropylidene-5′-oxo-5′-deoxyuridine under conditions of acid catalysis ($SnCl_4$). The intermediate dipyrryl–uridinyl methane was then subject to macrocyclization (BF_3 etherate) in the presence of 2′,3′-*O*-isopropylidene-5′-oxo-5′-deoxy-uridine or 4-fluorobenzaldehyde, followed by DDQ-mediated oxidation to produce the corresponding 2′,3′-*O*-isopropyliden-protected porphyrins, 5,10,15,20-tetrakisuridinylporphyrin **10a**, and 5,15-bisuridyl-10,20-(*p*-fluorophenyl)porphyrin **11a**, respectively.[12] Deprotection with aqueous trifluoroacetic acid at ambient temperature followed by neutralization with ammonia then yielded the free hydroxyl products **10b** and **11b** in nearly quantitative yield (Scheme 5).

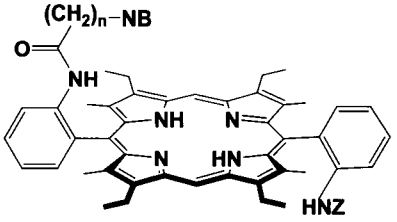
Yet another approach to generating nucleobase–porphyrin conjugates was pursued by Hisatome et al. These workers targeted the connection of a nucleobase moiety to a porphyrin via amide linkages, relying in particular on the reaction of an aminophenylporphyrin with an activated form of a nucleobase-bearing alkanoic acid.[13,14] Unfortunately, initial efforts to affect the coupling of *anti* 5,15-bis(2-aminophenyl)porphyrin derivatives with nucleobase–alkanoic acids using standard coupling agents such as *N,N′*-dicyclohexylcarbodiimide and *N,N′*-carbonyldiimidazole under a variety of conditions gave only small amounts of the expected amide or no product at all. On the other hand, very good results were obtained in the presence of ethoxycarbonyl chloride. However, in this case, only the corresponding singly functionalized, mixed monoamide monocarbonate derivatives **12** could be separated from the reaction mixture. In light of this finding, a decision was made to protect one of the two amino groups present in the starting bis-aminophenyl porphyrin with a benzyloxycarbonyl group.[13] Subsequent coupling with a nucleobase alkanoic acid afforded monofunctionalized compounds of general structure **13**.[13,14] The adenine and thymine derivatives **14** bearing free amino group were then prepared by removing the carbobenzyloxy groups from the corresponding precursors **13**. Coupling of thymine butanoic acid to **14a-A** then led to the *anti* adenine-thymine derivative **15**. The corresponding *syn* atropoisomer **18** was also prepared[13] using the same synthetic strategy, although in this instance the "intermediates" consisted of **16** and **17**, rather than **13** and **14** (Scheme 6).

As implied above, one of the prime motivations for generating porphyrinyl nucleoside, nucleotides, and oligonucleosides conjugates stemmed from a desire to produce water soluble, site-specific DNA binding and modifying systems.[15] While systems involving species containing oligonucleotides lie outside the scope of this review, it is worth noting that the covalent attachment of a photosensitizing porphyrin unit to an antisense oligonucleotide produced a system that allowed for the sequence-specific photomodification of target DNA after activation of the porphyrin subunit by photoillumination.[16]



10a R = R_1, R′-R′ = $C(CH_3)_2$ **11a** R = R_2, R′-R′ = $C(CH_3)_2$
10b R = R_1, R′ = H **11b** R = R_2, R′ = H

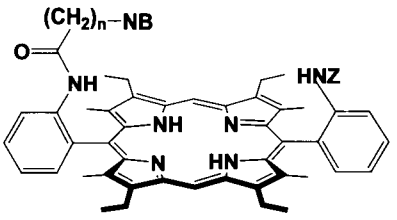
Scheme 5



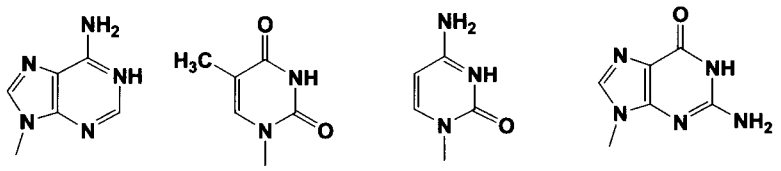
12 NB = A, T Z = $CO_2C_2H_5$
13 NB = A, T, C, G Z = $CO_2CH_2C_6H_5$
14 NB = A, T Z = H

for 12-14 a n = 3
b n = 4

15 NB = A n = 3 Z = $CO(CH_2)_3$-T



16 NB = A, T Z = $CO_2CH_2C_6H_5$
17 NB = A, T Z = H
18 NB = A n = 3 Z = $CO(CH_2)_3$-T



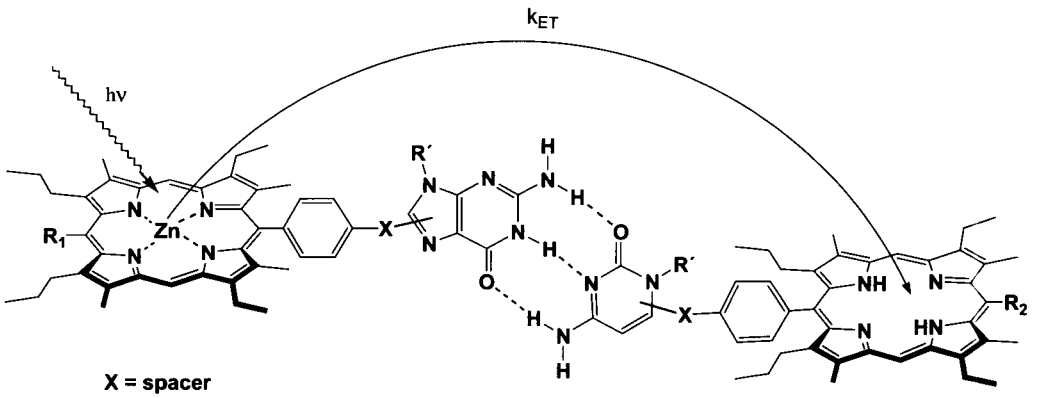
A T C G

Scheme 6

NB-substituted porphyrins as a model for electron transfer system

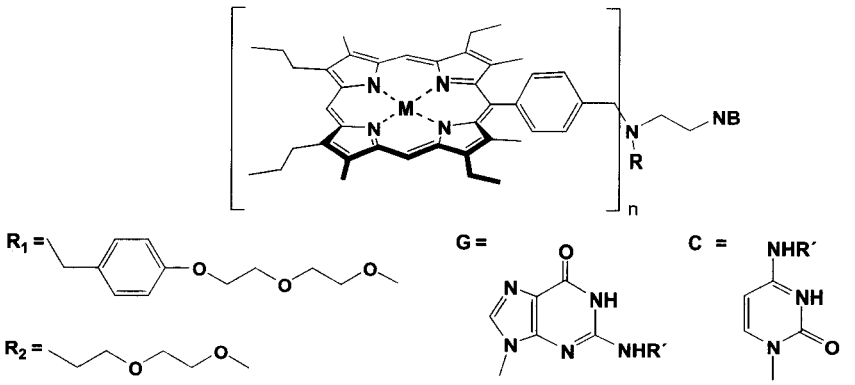
One of the more intriguing questions in the area of electron- and energy-transfer chemistry involves how covalent pathways differ from noncovalent ones in terms of mediating both short- and long-range transfer processes. Motivated by a desire to understand these issues, as well as to develop cytosine and guanine systems as new biomimetic modules for use in supramolecular ensemble construction, Sessler and coworkers prepared a series of noncovalent electron- and energy-transfer systems that contained porphyrinyl–nucleosides as key components.[17–20] These systems, which can be considered as containing a hydrogen-bonded assembled photon antenna, relied on Watson–Crick nucleobase pairing interactions between donors and aceptors attached to functionalized guanosine and cytosine subunits.

In one early system, guanosine- and cytidine-bearing methaleted (ZnP) and nonmethaleted (H_2P) porphyrins were used to construct an ensemble where photoexcitation led to energy transfer from the metalloporphyrin (characterized by a higher energy excited state) to a free-base porphyrin moiety (characterized by a lower



X = spacer

Fig. 1 Schematic representation of the hydrogen-bonded dimers **20–21** and subsequent energy transfer from ZnP **20** to H_2P **21**.



19a n = 1, R = R_1, NB = G, R′ = COPh, M = 2H
20a n = 1, R = R_1, NB = G, R′ = H, M = 2H
21a n = 1, R = R_1, NB = G, R′ = H, M = Zn

19b n = 1, R = R_2, NB = C, R′ = CPh_3, M = 2H
20b n = 1, R = R_2, NB = C, R′ = H, M = 2H
21b n = 1, R = R_2, NB = C, R′ = H, M = Zn

19c n = 2, NB = C, R′ = CPh_3, M = 2H
20c n = 2, NB = C, R′ = H, M = 2H
21c n = 2, NB = C, R′ = H, M = Zn

Scheme 7

energy excited state) in accord with Fig. 1. In this case, the requisite porphyrin–nucleobase conjugates **20** and **21** were prepared by alkylation of a nucleobase derivative bearing an amine "tail" with an electrophilic porphyrin. This provided the protected intermediates **19**. Subsequent deprotection afforded the free-base porphyrins **20** (H_2P), whose metalation gave the Zn(II) porphyrin analogues **21** (ZnP)[17,21] (Scheme 7).

The photophysical characteristics of the porphyrin–nucleobase conjugates **20** and **21** and the supramolecular dimers formed as a result of their undergoing Watson–Crick-type association (Fig. 1) were then determined by studying their spectroscopic properties in ethanol-free $CHCl_3$. Here, both standard static fluorescence and time-resolved methods were used. The latter proved particularly revealing. For instance, when solutions containing roughly equimolar concentrations of both H_2P **20** and ZnP **21** were subject to laser excitation at 570 nm, the excited ZnP singlet state (ZnP*) was formed as the dominant product (ca. 85%). Emission from this species, essentially free of fluorescence signals arising from the corresponding free-base excited singlet, H_2P*, could be monitored at 600 nm. Thus, varying the concentration of H_2P and studying the emission behavior of the whole system allowed the monitoring of the dimeric ensemble formation and energy transfer dynamics within it as well. At low concentrations of H_2P relative to ZnP, the fluorescence (as monitored at 600 nm) decayed with a lifetime τ_1 of $\sim$1.5 nsec, a value that corresponds to the singlet state lifetime of a simple, monomeric (i.e., noncomplexed) photoexcited zinc(II) porphyrin (i.e., ZnP*). As the relative concentration of H_2P increased, however, the decay profiles became increasingly dual exponential in character because of the appearance of a faster decaying component that was observed in addition to the longer ca. 1.5-nsec lifetime seen in the absence of H_2P. This shorter lifetime, τ_2 (ca. 0.6–0.8 nsec), was assigned to a Zn subunit held within a base-tethered dimer that was undergoing deactivation as a result of the intra-ensemble energy transfer occurring between ZnP* and H_2P. Energy transfer was observed also from the triplet excited state of the ZnP subunit, providing an ancillary set of data that could confirm the formation of the Watson–Crick associated ensemble. In fact, the extent of association could be calculated in quantitative terms from the observed singlet and triplet lifetimes. The resulting values, given in Table 1, confirmed the expected impression that the extent of nucleobase–nucleobase association increases in the order guanine–guanine<cytosine–cytosine<cytosine–guanine.[17]

In an effort to produce a model for electron-, as opposed to energy-, transfer, an ensemble involving the ZnP-guanosine derivative **21a** and the quinone cytosine derivative **22** was constructed via noncovalent self-as-

Table 1 Rate constants for energy transfer (k_{ss}) within the nucleic acid base dimers and association constants (K_a)

Porphyrins	τ_1 (nsec)	τ_2 (nsec)	$k_{ss}/10^8$ (sec^{-1})	K_a (M^{-1})	Reference
20c–21b	1.60	0.77	6.8	48	[17]
20a–21a	4.41	0.82	5.1	24	[17]
20b–21a	1.47	0.87	4.7	225	[17]
21a–22	1.50	0.94	4.2	1290	[18]
25–26	1.80	0.74	8.0	8990	[19]
24a–25	1.60	0.70	8.1	22,000	[20]

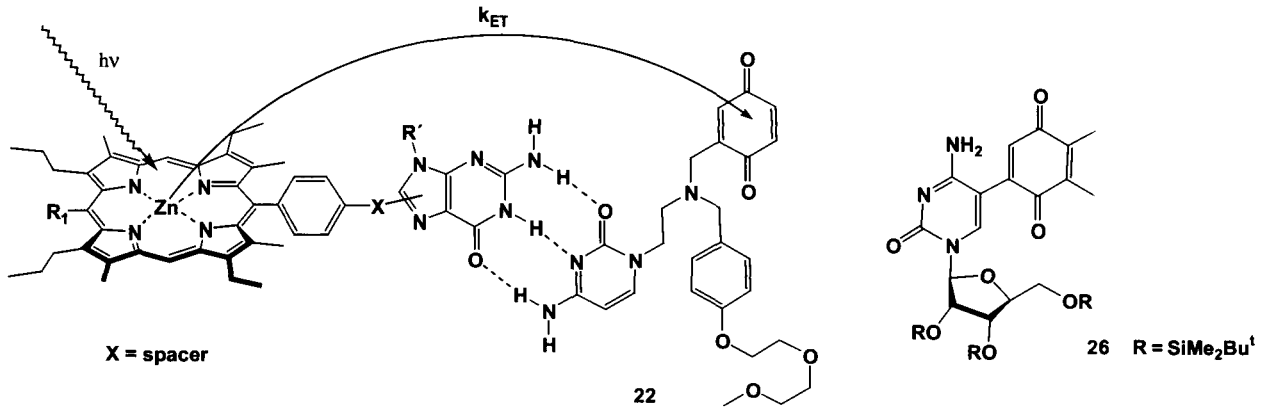


Fig. 2 Schematic representation of the porphyrin guanosine and quinone cytosine ensembles and subsequent energy transfer from porphyrin to quinone parts.

sembly (Fig. 2). In CH_2Cl_2, fluorescence from the porphyrin subunit in **21a** was increasingly quenched upon the addition of concentrations of **22** (to a limit of ca. 35% total quenching at the highest concentrations used). By contrast, no fluorescence changes were observed when similar experiments were carried out in the presence of methanol or when the exocyclic amine groups of **21a** were blocked. Time-resolved fluorescence studies of an equimolar mixture of porphyrinyl–guanosine **21a** and quinone cytosine **22** displayed a biphasic profile that was almost identical to that displayed by the self-assembled dimer of **20** and **21** (Table 1).[18] Such an observation was consistent with rapid intra-ensemble electron transfer in the case of the porphyrin–quinone pseudodimer formed from **21a** and **22**.

While important in terms of demonstrating the utility of Watson–Crick base pairing as a means of constructing noncovalent electron transfer model systems, the above first-generation system suffered from considerable conformation flexibility. Such flexibility was thought to limit the extent to which quenching could be achieved within the self-assembled pseudodimer formed from **21a** and **22** and, as such, complicated interpretation of the observed electron transfer process. For instance, quenching arising from inter- or intracomplex diffusional encounters between the donor and acceptor, rather than through a

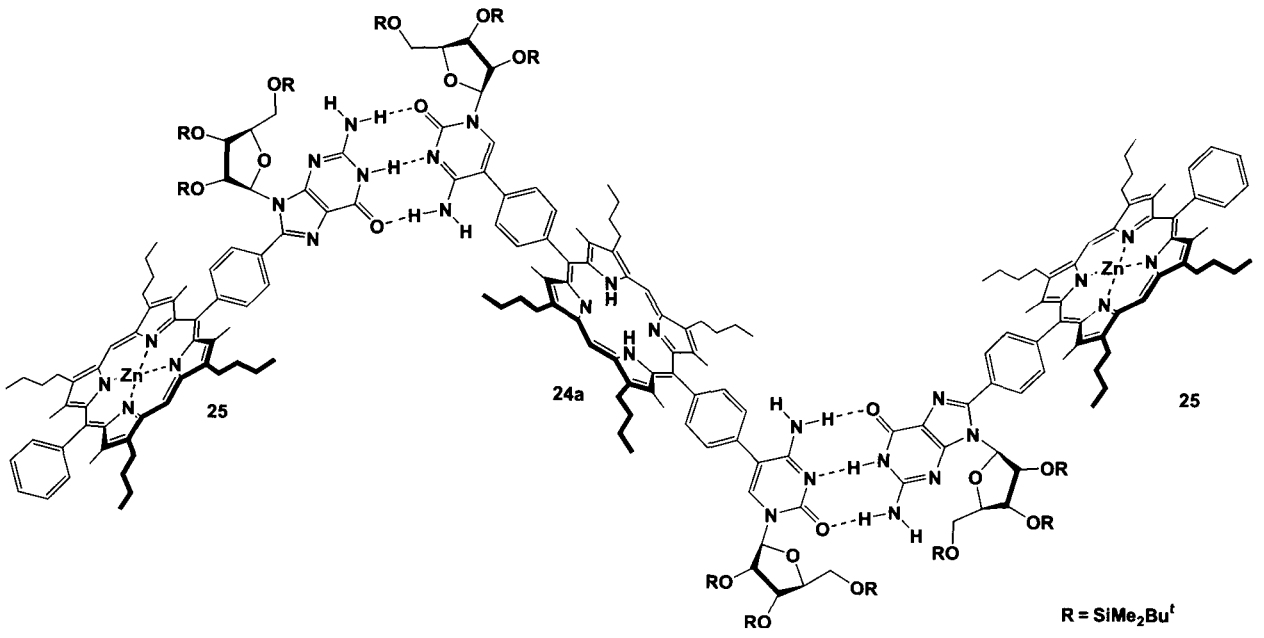
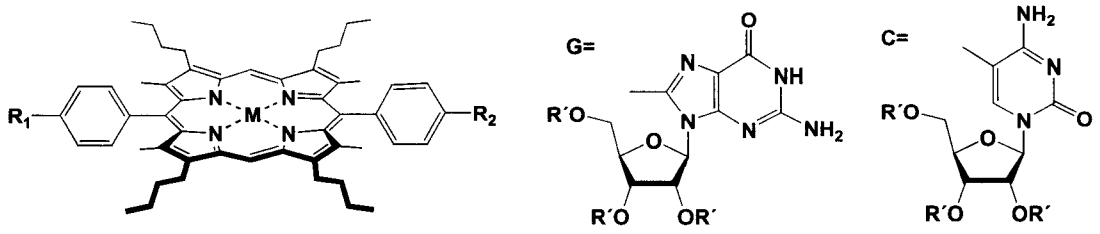


Fig. 3 Porphyrinic array wherein two ZnP **25** and one H_2P **24a** are held together in a trimeric conformation.



23 $R_1 = H$ $R_2 = G$, M = 2H
24a $R_1 = H$ $R_2 = C$, M = 2H
24b $R_1 = R_2 = C$, M = 2H

R′ = $SiMe_2Bu^t$

25 $R_1 = H$ $R_2 = G$, M = Zn

Scheme 8

hydrogen bond process, could be invoked to account for the observed quenching and presumed electron transfer events. Thus more rigid, second-generation, model systems (Fig. 3) were developed.[19,20] In particular, new porphyrin–nucleobase monomers were prepared from *O*- and *N*-protected guanosine- and cytidine-substituted benzaldehyde intermediates using the MacDonald–Chang porphyrin synthesis.[22] Subsequent deprotection and selective protection of the condensation products gave rise to the guanosine (**23**) and cytidine (**24**) porphyrins. Derivative **23** was then converted into its corresponding zinc porphyrin **25** (Scheme 8).

The noncovalent donor–acceptor systems derived from porphyrin **25** and either quinone cytidine **26** (Fig. 2) or porphyrin cytidine **24** (Fig. 3) were then subject to analysis via time-resolved fluorescence spectroscopy. In accord with the earlier, first-generation systems described above, addition of **24** or **26** to a chloroform solution of **25** caused a primarily monoexponential fluorescence decay profile to become biphasic. The resulting decay profiles were then analyzed in terms of two exponential lifetimes, corresponding to two excited state components (Table 1). The fractional amplitude of the shorter-lived component, considered to reflect an intra-ensemble quenching process, increased as the concentration of the acceptor component, either **24** or **26**, was increased. The corresponding association constants, K_a, were significantly higher than those derived for the more flexible aggregate[18] formed between **21a** and **22** ($K_a = 8990 \pm 600\ M^{-1}$ for the complex formed between **25** and **26** vs. $1290 \pm 230\ M^{-1}$).[19]

NB–Sapphyrin Derivative

Very different kinds of nucleobase–oligopyrrole conjugates may, in principle, be obtained when chromophores other than porphyrins are used. This is because other oligopyrrolic macrocycles are often characterized by properties that are very different than those of porphyrin. While a large number of porphyrin-like oligopyrrolic macrocycles now exist, the fact remains that, with the exception of porphyrin itself, most work in the NB–OPMC area has been largely limited to the use of sapphyrin. Sapphyrins are a class of pentapyrrolic expanded porphyrin that were first reported by Woodward (for a review, see Ref. [23]). Characterized by two relatively basic "pyridine-like" pyrrolic centers, sapphyrins are readily protonated and, in marked contradistinction to porphyrins, form complexes with a wide variety of anions, including in particular phosphates.[24,25] In fact, sapphyrin derivatives have been used to effect both nucleoside[26] and nucleotide[27] recognition. Furthermore, a silica-bound sapphyrin proved useful as a solid support for the HPLC separation of monomeric and short oligomeric nucleotides[28] at pH 7, while certain water-soluble sapphyrins were found to bind DNA in aqueous solution at physiological pH.[29] A variety of spectroscopic studies as well as X-ray diffraction analysis established that phosphate-type species are bound to the

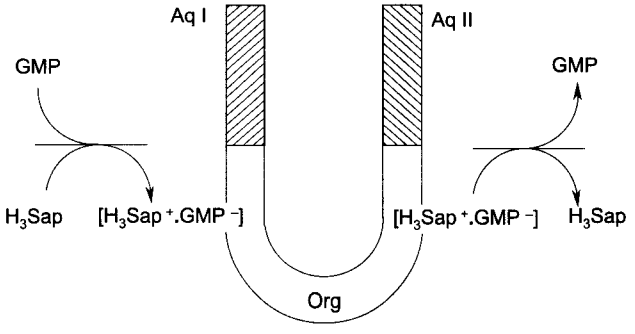
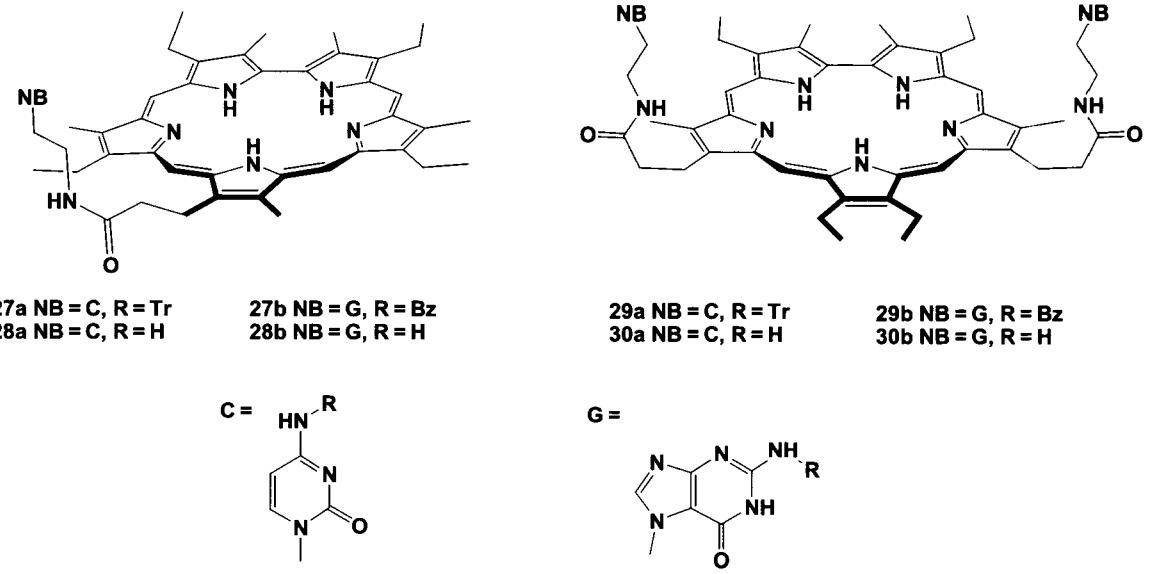


Fig. 4 Schematic representation of GMP transport effected under synport conditions using sapphyrin as the hydrophobic phosphate-binding carrier.



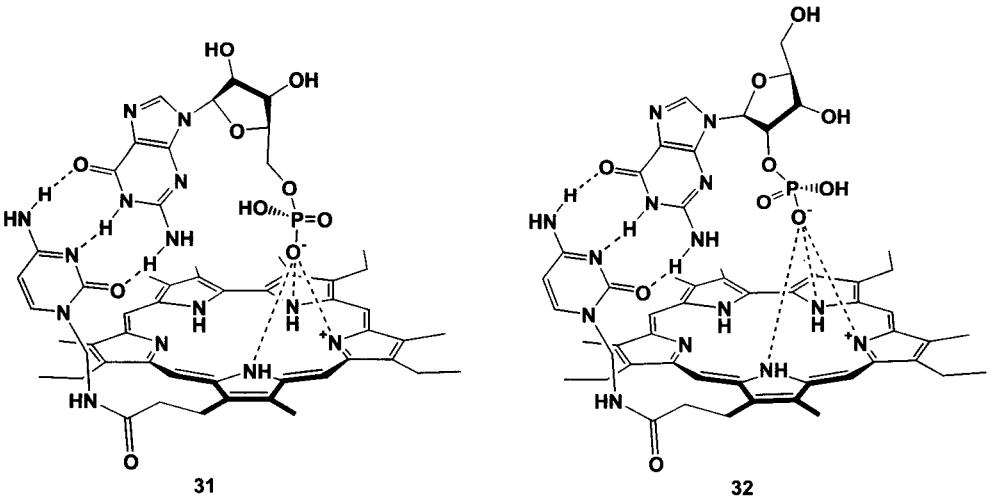
Scheme 9

protonated sapphyrin core via a "helicopter-like" set of hydrogen bonding interactions between the phosphate oxyanion and the pyrrolic hydrogens of the protonated macrocycle.[24]

Although unique, the phosphate binding capability of the sapphyrins is rather nonselective. For instance, the use of sapphyrin to affect the transport of nucleotide monophosphate through a model membrane (Fig. 4) consisting of an initial aqueous phase (Aq. I), an organic barrier, and a second, receiving phase (Aq. II) revealed that it is a very efficient but nonselective carrier at pH<4 (where it exists in its diprotonated form).[30] Such findings provided an incentive for Sessler and coworkers to develop more selective systems through the construction of NB–sapphyrin conjugates. These workers considered it likely that the combination of a NB (to recognize the complementary Watson–Crick nucleobase) and a sapphyrin (to effect phosphate binding) would lead to receptors with a unique ability to bind and transport nucleotides selectively. To the extent such species could be generated and might have an important application in the area of drug delivery; they could help affect the into-cell transport of nucleotide-based antiviral agents or, even in the limit, various antisense oligonucleotides. Important, however, as are these potential end-use applications, the initial goal was less ambitious; it was to find a sapphyrin-based nucleotide carrier that would work at physiological pH. Here, it was hoped that the

Table 2 Initial nucleotide-5′-monophosphate transport rates for the carriers **28a** and **30a**

Carrier	Aq. I (pH)	Aq. II	k_T 5′-GMP (10^{-8} mol/cm^2 hr)	$k_{5'\text{-GMP}}/k_{5'\text{-CMP}}$	$k_{5'\text{-GMP}}/k_{5'\text{-AMP}}$
28a	6.15	H_2O	1.201	101.7	7.66
28a	6.70	H_2O	0.287	42.9	8.87
28a	7.05	H_2O	0.001	20.1	9.49
28a	6.15	10 mM NaOH	1.423	26.3	2.73
28a	6.70	10 mM NaOH	1.228	40.8	4.36
28a	7.05	10 mM NaOH	0.708	43.3	9.60
30a	6.15	H_2O	0.101	6.2	1.38
30a	7.05	10 mM NaOH	0.115	23.7	3.18
None	7.00	H_2O	$<10^{-5}$		
Sapphyrin	7.00	H_2O	$<10^{-5}$		
Sapphyrin	7.00	10 mM NaOH	$<10^{-5}$		



31 32

Fig. 5 Proposed structure for the complex formed between the monoprotonated form of carrier **28a** and monobasic 5′-GMP or 2′-GMP, respectively.

"extra" recognition "power" derived by linking a nucleobase recognition subunit to a sapphyrin core would suffice to allow nucleotide transport at or near neutral pH.

In light of the above design considerations, systems **28** and **30** were chosen as initial targets. They were prepared by coupling protected derivatives of 2-aminoethylcytosine and 2-aminoethylcytosine with activated forms of either a sapphyrin carboxylic acid or sapphyrin diacid. Following deprotection[31] of the resulting intermediates, compounds **27** and **29**, the nucleobase functionalized sapphyrins **28** and **30** were obtained in good yield (Scheme 9).

As hoped, the cytosine-bearing carriers **28a** and **30a** exhibited selective transport of guanosine monophosphate through a model membrane at near-neutral pH. Interestingly, in all cases, receptor **28a** displayed a higher selectivity for GMP than its congener **30a** (Table 2). Control experiments were performed using 3, 8, 12, 13, 17, 22-hexaethyl-2, 7, 18, 23-tetramethylsapphyrin (sapphyrin).[32]

The regioselectivity of carrier **28a** for the phosphate group in isomeric GMPs was also studied (Fig. 5). Here, the 2′-isomer of GMP was transported roughly 10 and 6 times faster than its 5′- and 3′-substituted congeners, respectively (Table 3). The association constants, K_a, of complex **31** and **32** were determined to be 8×10^3 and 2.2×10^4 M^{-1}, respectively, in methanol as judged from UV–VIS spectroscopic titrations.[31]

In the case of the guanine-bearing sapphyrins **28b** and **30b**, transport selectivity for cytosine monophosphate was observed. Compared to **28b**, the relative rates of carrier-induced, through-membrane transport were found to be smaller for the monosubstituted system, **28b**, than for doubly functionalized analogue, **30b** (Table 4).[31]

To build on the above success, a new solid phase was prepared. It was produced by attaching a cytosine-substituted sapphyrin containing a carboxylic group to

Table 3 Initial rates of GMP isomer transport for carriers **28a** and **30a**

Carrier	Aq. I (pH)	Aq. II	k_T 2′-GMP (10^{-8} mol/cm^2 hr)	$k_{2'\text{-GMP}}/k_{5'\text{-GMP}}$	$k_{2'\text{-GMP}}/k_{3'\text{-GMP}}$
28a	6.70	H_2O	0.767	9.70	7.30
28a	6.70	1 mM NaOH	2.989	9.55	5.30
28a	7.00	H_2O	0.594	11.06	7.82
28a	7.20	H_2O	0.421	$>10^2$	$>10^2$
28a	7.35	H_2O	0.352	$>10^2$	$>10^2$
30a	6.70	1 mM NaOH	0.104	3.33	2.67
None	7.00	H_2O	$<10^{-5}$		
Sapphyrin	7.00	H_2O	2×10^{-5}		

Table 4 Initial nucleotide-5′-monophosphate transport rates for the carriers **28b** and **30b**

Carrier	Aq. I (pH)	Aq. II	k_T 5′-CMP (10^{-8} mol/cm^2 hr)	$k_{5'\text{-CMP}}/k_{5'\text{-GMP}}$	$k_{5'\text{-GMP}}/k_{5'\text{-AMP}}$
28b	6.70	H_2O	0.129	8.96	3.15
28b	6.70	1 mM NaOH	0.541	9.17	3.15
30b	6.70	1 mM NaOH	0.147	17.5	10.5

aminopropyl silica gel via an amide bound (Fig. 6). An HPLC column, generated by packing this stationary phase, separated guanosine 5′-mono-, di-, and triphosphate effectively from a mixture of the mono-, di-, and triphosphates of cytidine, uridine, adenosine, and guanosine under isocratic condition at pH 7. In addition, by using this column, it was found that all nucleotides monophosphate could be separated from one another readily under similar separation conditions.[33] Control experiments, affected using a column packed with a nucleoside-free sapphyrin-substituted silica gel,[28] allowed for the separation between mono-, di-, and triphosphates. However, no selectivity was observed. In other words, separation between the different mono-, di-, and trinucleotides present was not affected.[33]

NB–Phthalocyanine Derivatives

Although not expanded porphyrins, phthalocyanines represent a different kind of tetrapyrrole and one of the few aromatic chromophores, other than porphyrin and sapphyrin, to be conjugated to a nucleobase. While in principle the rich chemistry of the phthalocyanines would lead one to consider that a variety of nucleobase-containing conjugates would have been prepared to date, in fact, only the adenine-containing zinc(II) phthalocyanines **33** have so far been reported in the literature. These conjugates were prepared using a procedure analogous to that used to prepare the porphyrin–nucleobase systems **5** and **7**. Specifically, treatment of tetrahydroxyphthalocyanine with an excess of 9-(2-bromoethyl) adenine and K_2CO_3 gave the adenine–phthalocyanine conjugate **33a**. The unsymmetrical analogues **33b** and **33c** were prepared by using only one equiv. of the appropriate 9-(bromoalkyl)adenine, followed by the addition of excess 1-bromopentane[34] (Scheme 10).

The base-pairing effects of **33a** were studied by carrying out a fluorescence-type spectroscopic titration that consisted of adding a thymine-modified 9,10-anthraquione and monitoring the resulting decrease in emission intensity. Interestingly, in DMF/THF (9:1), the fluorescence quenching of **33b** and **33c** was modest with the extent of quenching being almost the same as was observed in experiments involving the use of 9,10-anthraquinone as a control.[34] While not constituting an absolute proof, such findings argue against the adenine-substituted conjugates **33** being able to interact strongly with species containing their Watson–Crick complement, thymine, and, as a consequence, an inability to form hydrogen bond-tethered self-assembled ensembles. If this conclusion is correct, it underscores the benefit of using cytosine–guanine interactions as supramolecular motifs as originally put forward by Sessler and coworkers.

NB–Calix[4]pyrrole Derivatives

While traditionally the term "oligopyrrole macrocycles" refers to aromatic systems, such as porphyrin, phthalo-

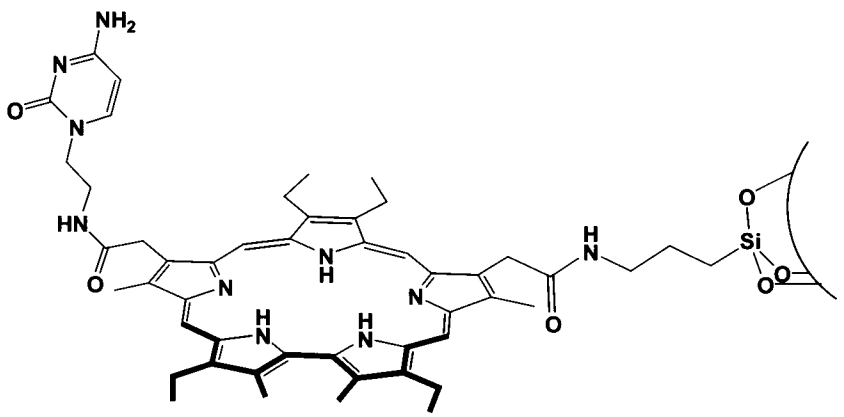
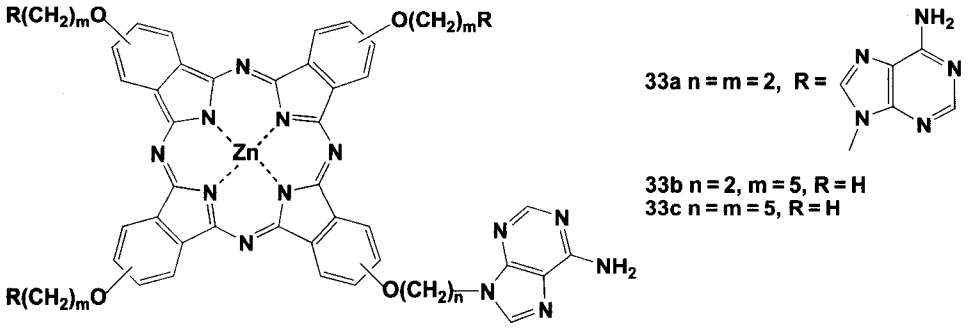


Fig. 6 Cytosine–sapphyrin derivative immobilized on 3-aminopropyl-modified silica.



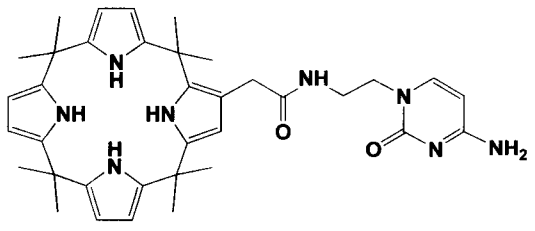
Scheme 10

cyanine, and expanded porphyrins, in recent years increasing attention has been devoted to the development and study of nonconjugated oligopyrrolic macrocycles.[35–38] One of the best studied of such systems is calix[4]-pyrrole,[35,36,39] a nonaromatic tetrapyrrolic product produced by the condensation of pyrrole with acetone. This system, whose synthetic origins can be traced back to Baeyer,[40] was "rediscovered" as a neutral anion-binding agent by Gale and coworkers in 1996.[39] This rediscovery led, inter alia, to considerations that calix[4]-pyrroles could provide the basis for easier-to-make analogues of the sapphyrin-based nucleotide binding and transporting agents described above.[41] Toward this end, receptors **34** and **35** were prepared by condensing the corresponding calix[4]pyrrolecarboxylic acid with 2-aminoethylcytosine.[42] Again, the question was whether the combination of an anion binding subunit with a nucleobase binding site would allow for selective nucleotide recognition and transport, with the added wrinkle being what would be the effect of using a weaker anion binding site (Scheme 11). (As they are neutral receptors, calix[4]-pyrroles show much weaker phosphate anion binding affinities than do the protonated, and correspondingly charged, sapphyrins.[41])

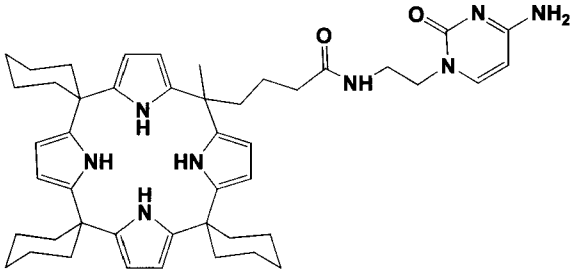
In order to test the efficacy and selectivity of the calix[4]pyrrole-derived systems **34** and **35**, they were studied as nucleotide carriers, in analogy to what was done with the sapphyrin systems (vide supra), and as the key components in membrane-based ion selective electrodes (ISEs) (see Tables 5–7 and ensuing discussion).

Ion-Selective Electrode Studies

An important motivation for studying systems **34** and **35** as carrier-based ion-selective electrodes (ISE) is that analysis of such systems provides another means of testing whether a given receptor displays selectivity for a targeted analyte.[27,43–48] As true for bulk membrane transport studies (cf. Fig. 4), this method can provide insight into recognition events that take place at an aqueous–organic interface. However, it does not directly monitor binding (and/or release) per se. Rather, what is studied is the change in membrane potential observed on exposure of a liquid/polymer membrane electrode to solutions of various putative analytes. Read-out parameters thus include response (total emf change engendered by a given concentration of analyte), sensitivity (change in emf as a function of analyte concentration), and selectivity, often expressed in relative terms as a selectivity coefficient $K_{I/J}^{Pot}$ or selectivity factor $k_{I/J}^{Psel}$, where I and J represent the two competing analytes in question,



34



35

Scheme 11

Table 5 Results of transport experiments carried out with β-(**34**) and *meso*-(**35**) cytosine-functionalized calix[4]pyrroles

Carrier	k_T 5′-GMP (10^{-11} mol/cm^2 hr)	$k_{5'\text{-GMP}}/k_{5'\text{-CMP}}$	$k_{5'\text{-GMP}}/k_{5'\text{-AMP}}$
34	9.8	0.2	1.8
35	93	7.8	1.9

These transport studies were carried out using the set-up shown in Fig. 4 and in analogy to those discussed previously in the case of the sapphyrin-derived systems. The concentration of the carriers in the organic phase was 0.1 mM, the pH of the initial and receiving aqueous phases were 6.0 and 12.5, respectively, and the concentration of tetrabutylammonium perchlorate, added as a charge neutralizing cocarrier to the organic phase, was 0.1 mM.

and the linear range over which the response, Nernstian or otherwise, is seen.

The original benchmark for nucleotide sensors was an electrode for 5′-adenosine monophosphate (5′-AMP) that was produced by Papastathopoulos and Rechnitz.[49] This electrode consisted of a layer of suspended 5′-adenylic acid deaminase (AMP deaminase) in conjunction with an ammonia gas-sensing membrane electrode. The substrate was selectively deaminated by this enzyme to produce inosine 5′-monophosphate (5′-IMP) and NH_3 in stoichiometric quantities. Monitoring the ammonia thus indirectly measured the 5′-AMP concentration present in the sample.

Because the above method represents an indirect determination, efforts have been devoted in recent years to developing more biomimetic approaches. Complementary base pairing, directed multisite hydrogen bonding interactions, specific (π–π) stacking effects, and generalized electrostatic interactions, individually or in concert, are thought to contribute to the exquisitely sensitive nucleotide recognition observed in the biological systems. Incorporating one or more of these recognition motifs into appropriately designed synthetic receptors could culminate in the production of highly specific ISEs more suited for use in various "real world" bioanalytical applications. Whether this goal will be fully realized remains to be seen. However, its pursuit has stimulated the evolutionary development of several elegant nucleotide-targeting ISEs in recent years, including (as will be discussed below) those based on calix[4]pyrrole–nucleobase conjugates.

An electrode derived from a cytosine-pendant triamine was the first ISE system to affect the potentiometric discrimination between guanine and adenine nucleotides by using complementary base-pairing and phosphate–ammonium electrostatic binding interactions within the electrode.[47] However, the multiple protonation equilibria that were a consequence of using polyamines as the electrostatic binding motif likely complicated the response mechanism and were also most probably responsible for the non-Nernstian emf slopes observed in these systems (5′-GMP: -10 mV/decade; 5×10^{-4}–10^{-2} M).

Recognizing the design limitations, it was thought that better phosphate recognition motifs would lead to improved sensors. In a previous portion of this review, it was demonstrated how this approach led to successful sapphyrin-based nucleotide carriers (e.g., systems **27a** and **28a**)[31,32,50] and, as implied above, efforts were made to generalize this success by using calix[4]pyrroles as the anion binding cores. Thus because they were of inherent interest in their own right and because they would provide important controls for the newer calix[4]pyrrole-based systems, ISEs containing nucleobase-functionalized sapphyrins were prepared. Specifically, the guanine-bearing sapphyrins (**28b** and **30b**) were incorporated into poly (vinyl chloride) (PVC) membranes plastisized by *o*-nitrophenyl octyl ether (*o*-NPOE), and the potential response as a function of nucleotide concentration was measured using a standard Hg|Hg_2Cl_2|3 M KCl‖0.1 M HEPES–NaOH, pH 6.6‖sample|modified PVC membrane|0.1 M KCl|AgCl|Ag cell assembly. Similar studies were carried

Table 6 Potentiometric characteristics of PVC membranes based on mono-(**28b**) and bisguanine-substituted (**30b**) sapphyrins toward nucleotides

	28b		**30b**[a]		**30b**[b]	
Nucleotide	**Sensitivity (mV/decade)**	**Linear range (M)**	**Sensitivity (mV/decade)**	**Linear range (M)**	**Sensitivity (mV/decade)**	**Linear range (M)**
5′-AMP	−32	10^{-6}–10^{-5}	−42	10^{-6}–10^{-5}	−30	10^{-6}–10^{-5}
5′-CMP	−17	10^{-5}–10^{-4}	−59	10^{-6}–10^{-4}	−20	10^{-6}–10^{-3}
5′-GMP	c	c	−25	10^{-5}–10^{-3}	−58	10^{-6}–10^{-4}

[a,b]The results were obtained with PVC-membranes derived from **30b** soaked in solutions of the primary analytes for 20 min (for a) and overnight (for b), respectively.
[c]No response.

out using the cytosine-substituted calix[4]pyrrole conjugates **34** and **35**.

Guanine-substituted sapphyrins

The ditopic receptor **28b** was designed to allow for specific cytosine-base recognition as well as for more general phosphate binding recognition. Its bis(guanine)-bearing analogue, sapphyrin **30b**, a potential tritopic receptor, was thought to allow for the formation of complex-derived "triple helix" such as G–C–G motifs, wherein the bound nucleobase C-subunit would be the beneficiary of two kinds of hydrogen bonding interactions, involving both Watson–Crick and Hoogsteen recognition patterns.

At pH 6.6, sapphyrin exists as a monoprotonated, singly charged entity, whereas nucleotide monophosphates are largely dianionic. The sensitivity expected for experimental electrodes at neutral pH should thus be equal to −29 mV/decade (theoretical Nernstian value for a divalent anion). The results obtained with systems **28b** and **30b** (Table 6) revealed that the doubly functionalized guanine derivative, **30b**, provided a more effective electrode system than its mono-substituted analogue. It should be noted that the sensitivity of the PVC-membranes derived from the bisguanine receptor **30b** was dependent on the time of contact between the membrane and the nucleotide (analyte) solution, as well as the nature of the measured nucleotide (Table 6).

Cytosine-substituted calix[4]pyrroles

In predicative work involving unfunctionalized calix[4]-pyrroles,[45] it was demonstrated that simple, substituent-free calix[4]pyrroles display potentiometric selectivity to a range of anionic analytes, including phosphates, when incorporated into PVC-(*o*-NPOE) membranes and tested as ISEs. The question then became whether appending a cytosine "tail" onto the calixpyrrole skeleton would lead to the generation of nucleotide-specific ISEs and, to the extent this proved true, whether or not the choice of linkage (*meso*- vs. *β*-pyrrolic) would effect the response selectivity.[42]

Before analyzing the "tailed" systems **34** and **35**, their unfunctionalized analogues, i.e., octamethyl and cyclohexyl-substituted systems, were tested as ISE sensor elements. In neither case was evidence of pH-dependent behavior seen, at least at or near neutral pH. On the other hand, an inherent selectivity for 5′-AMP<5′-GMP≈5′-CMP< 5′-UMP≈5′-TMP was observed at pH 6.6 (a value chosen to ensure a significant concentration of the dianionic forms of the nucleotides under investigation) for both systems, as judged from the extent of the anionic (negative) potentiometric response. By contrast, a slight selectivity for 5′-GMP and 5′-CMP was seen in the case of "control" electrodes made up from the hydrophobic cation, tridodecylmethylammonium chloride (TDDMACl); here, 5′-AMP (0.00)<5′-UMP (0.07)<5′-CMP (0.24)<5′-GMP (0.76), where the values in parentheses refer to the selectivity coefficients (log $K^{Pot}_{5'\text{-AMP/5-XMP}}$).

Taken together, these findings are consistent with the conclusion that unfunctionalized calixpyrroles mediate their observed ISE response for nucleotides by acting more as specific, nucleobase-dependent molecular recognition elements than as pure anion extractants, as is known to be true for membranes made up from TDDMACl. Support for this conclusion comes from the observation that a greater response is observed in the case of the more hydrophobic cyclohexyl-substituted system than in the octamethyl system and that the selectivity pattern correlates with the number of accessible hydrogen bond acceptor elements (i.e., carbonyl groups) present in the nucleobase portion of the mononucleotides being studied [i.e.,5′-UMP≈5′-TM (two carbonyls)>5′ GMP≈5′-CMP (one carbonyl)>5′-AMP (no carbonyls)]. This latter observation also rules out a response process that is dominated by direct phosphate-calixpyrrole "anion chelation"; rather, it supports the conclusion that, under the interfacial conditions of the ISE experiment, the strength and specificity of the nucleobase–calix[4]pyrrole NH interactions dominate the selectivity, even if it is the presence of the negatively charged phosphate groups that leads to the actual observation of an anionic potentiometric response.

Transport experiment demonstrated that the *meso*-substituted cytosine calix[4]pyrrole conjugate, but not its *β*-pyrrole linked congener, is capable of acting as a ditopic receptor, binding concurrently both the phosphate anion and nucleobase portions of 5′-GMP to the calixpyrrole core and cytosine "tails" of the molecule, respectively. These findings were rationalized in terms only the former is capable of binding concurrently both the phosphate and nucleobase portions and provided a backdrop against which the corresponding ISE studies could be considered. In particular, the availability of two cytosine-substituted calix[4]pyrroles, bearing the appended cytosine on either a *β*-pyrrolic (**34**) or *meso* (**35**) position, respectively, would allow the potential importance that ditopic binding interactions might play in terms of regulating a selective electrode response.

In the case of the transport studies, it proved necessary to add tetrabutylammonium perchlorate to neutralize the negative charge present in the calix[4]pyrrole–nucleotide complex that arises as a result of binding a phosphate-containing entity within a neutral receptor (cf. Table 5 and Ref. [42]). Thus a lipophilic electrically charged additive, such as TDDMACl, was also employed in the case of the PVC membranes derived from **33** and **35**. Addition of this

Table 7 Effect of pH on the potentiometric selectivity of PVC membranes based on β- and *meso*-substituted cytosine-functionalized calix[4]pyrroles **34** (log $k^{sel}_{5'\text{-UMP/5-XMP}}$) and **35** (log $k^{sel}_{5'\text{-UMP/5-XMP}}$). These membranes additionally contained 50 mol% TDDMACl

	34			35		
pH	**5′-AMP**	**5′-CMP**	**5′-GMP**	**5′-AMP**	**5′-CMP**	**5′-GMP**
6.6	−0.07	+0.97	+0.89	+1.53	0.00	+2.23
8.5	−0.20	+0.58	+0.10	−0.68	0.00	−1.23

additive served to highlight the effects of base-pairing, which proved more discernable in the case of **35** than **34**. Specifically, adding 50 mol% TDDMACl to the PVC membranes containing these two potentially ditopic receptors resulted in selectivity sequences that approximated those seen in the transport experiments (Table 7; pH 6.6). Unfortunately, the inherent selectivity for 5′-GMP seen in membranes containing just TDDMACl (vide supra) complicates analysis. Thus it is not possible at present to quantify in energetic terms the specific effect of the proposed cytosine–guanosine Watson–Crick binding interactions on the experimentally determined for-analyte selectivities. Nonetheless, there is little doubt that the effect is real, and in ISEs based on suitably designed receptors, can be quite substantial.

Model membrane transport studies provide a useful potential complement to ISE studies in terms of analyzing the specificity and mode of action of a particular receptor. The fact that the transport selectivity may be influenced strongly by the rate of receptor–substrate complex formation and dissociation, and hence the pH of both aqueous phases, made it important to compare the potentiometric selectivity of the PVC-membrane electrodes based on **34** and **35** when the contacting aqueous phase was held at pH 6.6 and at 8.5. The chosen pH should approximate the conditions of potentiometric measurement to transport model when the monitored nucleotides "enter in" and "release from" bulk membrane interface. As can be seen from Table 7, the selectivity sequence for the *meso*-linked conjugate **35** is more influenced by pH than its β-linked congener **34**. In fact, the potentiometric selectivity of the meso-substituted receptor, **35**, could be reversed by changing the pH from 6.6 to 8.5, as is evident upon inspection of the following sequences:

pH 6.6 5′-CMP (0.00)<5′-AMP (1.53)<5′-GMP (2.23)

pH 8.5 5′-GMP (−1.33)<5′-AMP (−0.68)<5′-CMP (0.00)

The sensitivity to pH is as would be expected for a system where ditopic interactions are thought to be playing an important role. For instance, under conditions of high pH, where the nucleotide species would exist as a dianion, rather than a monoanion, and the nucleobase itself could bear negative charge, the ancillary benefit that would accrue from the proposed Watson–Crick base pairing interactions would be reduced if not altogether eliminated. On the other hand, the potentiometric data at high pH could help explain the relative lack of for-GMP selectivity observed in the case of the transport experiments involving receptor **34**, where the receiving phase is also held at high pH; what is observed might simply reflect the inherent preference for cytosine binding that apparently pertains at high pH. To the extent this conclusion is correct, it underscores the benefit that can come from studying nucleobase–oligopyrrole conjugates under two rather different kinds of experimental conditions and via the use of two very different techniques. It is thus hoped that such comparisons will be carried out with other systems in the near future.

CONCLUSION

Nucleobase–macrocyclic oligopyrrole conjugates represent a new class of potential ditopic receptor. With the ability to bind the Watson–Crick complementary nucleoside or nucleotide via specific hydrogen bonding interactions and elicit some key, useful functional feature through the appended oligopyrrole, a range of chemical processes become possible that are without precedent in the case of simple, single-motif systems. These include the ability to affect selective recognition of appropriate oligonucleotide fragments, facilitate long-range energy- and electron-transfer processes, affect the specific transport and electrochemical sensing of individual nucleotides. Improvements in synthetic methodologies, coupled with an ever-increasing lexicon of available oligopyrrolic macrocycles, lead to the prediction that many new nucleobase–oligopyrrole conjugates will be constructed in the near future and that these systems will demonstrate utilities that range well beyond those that have already been described.

ACKNOWLEDGMENTS

Financial support from the National Institutes of Health (GM 58907 to J.L.S.), financial support from the Ministry

of Education of the Czech Republic Grant No. MSM 223400008, the Grant EU QLRT-2000-02360, and Grant Agency of the Czech Republic Nos. 301/98/K042, 203/03/0716, 203/02/0420, and 309/02/1193 are gratefully acknowledged.

REFERENCES

1. Kus, P.; Knerr, G.; Czuchajowski, L. First representatives of porphyrinylnucleosides. Tetrahedron Lett. **1990**, *31* (36), 5133–5134.
2. Czuchajowski, L.; Habdas, J.; Niedbala, H.; Wandrekar, V. Synthesis of porphyrinyl–nucleosides. J. Heterocycl. Chem. **1992**, *29* (2), 479–486.
3. Lipscomb, L.A.; Zhou, F.X.; Presnell, S.R.; Woo, R.J.; Peek, M.E.; Plaskon, R.R.; Williams, L.D. Structure of a DNA–porphyrin complex. Biochemistry **1996**, *35* (9), 2818–2823.
4. Maldoti, A.; Andreotti, L.; Molinari, A.; Borisov, S.; Vasilev, V. Photoinitiated catalysis in Nafion membranes containing palladium(II) *meso*-tetrakis(*N*-methyl-4-pyridyl)porphyrin and iron(III) *meso*-tetrakis(2,6-dichlorophenyl)porphyrin for O_2-mediated oxidations of alkenes. Chem. Eur. J. **2001**, *7* (16), 3564–3571.
5. Czuchajowski, L.; Habdas, J.; Niedbala, H.; Wandrekar, V. Porphyrinyl–uridines as the first water soluble porphyrinyl–nucleosides. Tetrahedron Lett. **1991**, *32* (51), 7511–7512.
6. Czuchajowski, L.; Niedbala, H. Synthesis and tumoricidal activity of water soluble porphyrinyl-thymidines and related porphyrins. Bioorg. Med. Chem. Lett. **1992**, *2* (12), 1645–1648.
7. Czuchajowski, L.; Palka, A.; Morra, M.; Wandrekar, V. Porphyrinyl-nucleosides containing fluorinated nucleobases. Tetrahedron Lett. **1993**, *34* (34), 5409–5412.
8. Malinovski, V.; Tumir, L.; Piantanida, I.; Zinic, M.; Schneider, H.-J. New porphyrin-nucleobase hybrid compounds and their interaction with nucleosides and nucleic acids. Eur. J. Org. Chem. **2002**, (22), 3785–3795.
9. Masiero, S.; Gottarelli, G.; Pieraccini, S. G-quartets as a self-assembled scaffold for circular porphyrin arrays. Chem. Commun. **2000**, (20), 1995–1996.
10. Li, H.; Czuchajowski, L. Compounds based on *meso*-tri(4-pyridyl)-*p*-acrylamidophenylporphyrin able to interact with DNA. J. Heterocycl. Chem. **1997**, *34* (3), 999–1003.
11. Goh, G.K.-M.; Czuchajowski, L. The synthesis of isomeric dithymidyl-phosphorus(V)-*meso*-tetraphenylporphyrins. J. Porphyr. Phthalocyanines **1997**, *1* (3), 281–285.
12. Cornia, M.; Binacchi, S.; Del Soldato, T.; Zanardi, F.; Casiraghi, G. Synthesis of novel porphyrin–uridine carbon–carbon conjugates. J. Org. Chem. **1995**, *60* (16), 4964–4965.
13. Hisatome, M.; Maruyama, N.; Ikeda, K.; Furutera, T.; Ishikawa, T.; Yamakawa, K. Synthesis and some spectroscopic properties of porphyrin derivatives connected with nucleobases (adenine, thymine, guanine and cytosine) by alkanamide chains. Chem. Pharm. Bull. **1996**, *44* (10), 1801–1811.
14. Hisatome, M.; Maruyama, N.; Furutera, T.; Ishikawa, T.; Yamakawa, K. Porphyrine coupled with nucleoside bases. Synthesis and characterization of adenine- and thymine-porphyrin derivatives. Chem. Lett. **1990**, (12), 2251–2254.
15. Li, H.; Czuchajowski, L. Synthesis and biomedical application of porphyrinyl nucleosides, nucleotides and oligonucleotides. Trends Heterocycl. Chem. **1999**, *6*, 57–77.
16. Seliger, H.; Knoller, H.; Ruck, A.; Heckelsmiller, K.; Steiner, R. Antisense oligonucleotide conjugates with photosensitizers—An update. Nucleosides Nucleotides **1998**, *17* (9–11), 2053–2061.
17. Harriman, A.; Magda, D.J.; Sessler, J.L. Photon antennae assembled by nucleic acid base pairing. J. Phys. Chem. **1991**, *95* (4), 1530–1532.
18. Harriman, A.; Kubo, Y.; Sessler, J.L. Molecular recognition via base pairing: Photoinduced electron transfer in hydrogen-bonded zinc porphyrin–benzoquinone conjugates. J. Am. Chem. Soc. **1992**, *114* (1), 388–390.
19. Sessler, J.L.; Wang, B.; Harriman, A. Long-range photoinduced electron transfer in an associated but noncovalently linked photosynthetic model system. J. Am. Chem. Soc. **1993**, *115* (22), 10418–10419.
20. Sessler, J.L.; Wang, B.; Harriman, A. Photoinduced energy transfer in associated but noncovalently linked photosynthetic model systems. J. Am. Chem. Soc. **1995**, *117* (2), 704–714.
21. Harriman, A.; Magda, D.J.; Sessler, J.L. Energy transfer across a hydrogen-bonded, cytosine-derived, zinc-free-base porphyrin conjugate. J. Chem. Soc., Chem. Commun. **1991**, (5), 345–347.
22. Chang, C.K.; Abdalmuhdi, I. Anthracene pillared cofacial diporphyrin. J. Org. Chem. **1983**, *48* (26), 5388–5390.
23. Sessler, J.L.; Davis, J.M. Sapphyrins: Versatile anion binding agents. Acc. Chem. Res. **2001**, *34* (12), 989–997.
24. Sessler, J.L.; Cyr, M.J.; Furuta, H.; Král, V.; Mody, T.; Morishima, T.; Shionoya, M.; Weghorn, S. Anion binding: A new direction in porphyrin-related research. Pure Appl. Chem. **1993**, *65* (3), 393–398.
25. Král, V.; Furuta, H.; Shreder, K.; Lynch, V.; Sessler, J.L. Protonated sapphyrins. Highly effective phos-

phate receptors. J. Am. Chem. Soc. **1996**, *118* (7), 1595–1607.
26. Záruba, K.; Tománková, Z.; Sýkora, D.; Charvátová, J.; Kavenová, I.; Bouř, P.; Matějka, P.; Volka, K.; Král, V. Interaction of porphyrin and sapphyrin macrocycles with nucleobases and nucleosides. Spectroscopic, quantum chemical and chromatographic investigation. Anal. Chim. Acta **2001**, *437*, 39–53.
27. Tohda, K.; Naganawa, R.; Lin, X.M.; Tange, M.; Umezawa, K.; Odashima, K.; Umezawa, Y.; Furuta, H.; Sessler, J.L. Liquid membrane electrodes for nucleotides based on sapphyrin, cytosine-pendant triamine and neutral cytosine derivative as sensory elements. Sens. Actuators, B **1993**, *14* (1-3), 669–672.
28. Iverson, B.L.; Thomas, R.E.; Král, V.; Sessler, J.L. Molecular recognition of anionic species by silica gel bound sapphyrin. J. Am. Chem. Soc. **1994**, *116* (6), 2663–2664.
29. Iverson, B.L.; Shreder, K.; Král, V.; Sansom, P.; Lynch, V.; Sessler, J.L. Interaction of sapphyrin with phosphorylated species of biological interest. J. Am. Chem. Soc. **1996**, *118* (7), 1608–1616.
30. Furuta, H.; Cyr, M.J.; Sessler, J.L. Phosphate anion binding: Enhanced transport of nucleotide monophosphates using a sapphyrin carrier. J. Am. Chem. Soc. **1991**, *113* (17), 6677–6678.
31. Král, V.; Sessler, J.L. Molecular recognition via base-pairing and phosphate chelation. Ditopic and tritopic sapphyrin-based receptors for the recognition and transport of nucleotide monophosphates. Tetrahedron **1995**, *51* (2), 539–554.
32. Král, V.; Sessler, J.L.; Furuta, H. Synthetic sapphyrin–cytosine conjugates: Carriers for selective nucleotide transport at neutral pH. J. Am. Chem. Soc. **1992**, *114* (22), 8704–8705.
33. Sessler, J.L.; Genge, J.W.; Král, V.; Iverson, B.L. Separation of mono-, di- and triphosphate nucleotides by cytosine substituted, silica-bound sapphyrin solid supports. Supramol. Chem. **1996**, *8*, 45–52.
34. Li, X.; Ng, D.K.P. Synthesis and spectroscopic properties of the first phthalocyanine-nucleobase conjugates. Tetrahedron Lett. **2001**, *42* (2), 305–309.
35. Gale, P.A.; Sessler, J.L.; Král, V. Calixpyrroles. Chem. Commun. **1998**, 1–8.
36. Gale, P.A.; Anzenbacher, P., Jr.; Sessler, J.L. Calixpyrroles II. Coord. Chem. Rev. **2001**, *222* (1), 57–102.
37. Sessler, J.L.; Zimmerman, R.S.; Bucher, C.; Král, V.; Andrioletti, B. Calixphyrins. Hybrid macrocycles at the structural crossroads between porphyrins and calixpyrroles. Pure Appl. Chem. **2001**, *73* (7), 1041–1057.
38. Sessler, J.L.; Maeda, H.; Mizuno, T.; Lynch, V.; Furuta, H. Quinoxaline-bridged porphyrinoids. J. Am. Chem. Soc. **2002**, *124* (45), 13474–13479.
39. Gale, P.A.; Sessler, J.L.; Král, V.; Lynch, V. Calix[4]pyrroles: Old yet new anion-binding agents. J. Am. Chem. Soc. **1996**, *118* (21), 5140–5141.
40. Baeyer, A. Über ein condensationsproduct von pyrrol mit aceton. Ber. Dtsch. Chem. Ges. **1886**, *19*, 2184–2185.
41. Allen, W.E.; Sessler, J.L. Anion carriers: New tools for crossing membranes. ChemTech **1999**, *29* (9), 16–24.
42. Sessler, J.L.; Král, V.; Shishkanova, T.V.; Gale, P.A. Cytosine substituted calix[4]pyrroles: Neutral receptors for 5′-guanosine monophosphate. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99* (8), 4848–4853.
43. Amemiya, S.; Buehlmann, P.; Tohda, K.; Umezawa, K. Hydrogen bond based recognition of nucleotides by neutral-carrier ion-selective electrodes. Anal. Chim. Acta **1997**, *341* (2–3), 129–139.
44. Buhlmann, P.; Amemiya, S.; Nishizawa, S.; Xiao, K.P.; Umezawa, Y. Hydrogen-bonding ionophores for inorganic anions and nucleotides and their application in chemical sensors. J. Incl. Phenom. Mol. Recognit. Chem. **1998**, *32* (2-3), 151–163.
45. Král, V.; Sessler, J.L.; Shishkanova, T.V.; Gale, P.A.; Volf, R. Molecular recognition at an organic–aqueous interface: Heterocalixarenes as anion binding agents in liquid polymeric membrane ion-selective electrodes. J. Am. Chem. Soc. **1999**, *121* (38), 8771–8775.
46. Odashima, K.; Naganawa, R.; Radecka, H.; Kataoka, M.; Kimura, E.; Koike, T.; Tohda, K.; Tange, M.; Furuta, H.; Sessler, J.L.; Yagi, K.; Umezawa, Y. Chemical sensing based on membrane potential change induced by host–guest complexation at a membrane surface. Supramol. Chem. **1994**, *4*, 101–113.
47. Tohda, K.; Tange, M.; Odashima, K.; Umezawa, K.; Furuta, H.; Sessler, J.L. Liquid membrane electrode for guanosine nucleotides using a cytosine-pendant triamine host as the sensory element. Anal. Chem. **1992**, *64* (8), 960–964.
48. Umezawa, Y.; Kataoka, M.; Takami, W.; Kimura, E.; Koike, T.; Nada, H. Potentiometric adenosine triphosphate polyanion sensor using a lipophilic macrocyclic polyamine liquid membrane. Anal. Chem. **1988**, *60* (21), 2392–2396.
49. Papastathopoulos, D.S.; Rechnitz, G.A. Highly selective enzyme electrode for 5-adenosine monophosphate. Anal. Chem. **1976**, *48* (6). 862–864.
50. Sessler, J.L.; Furuta, H.; Král, V. Phosphate anion chelation and base-pairing. Design of receptors and carriers for nucleotides and nucleotide analogues. Supramol. Chem. **1993**, *1* (3–4), 209–220.

Oil-Filled Nanocapsules

Royale S. Underhill
Defence Research and Development Canada—Atlantic, Dartmouth, Nova Scotia, Canada

INTRODUCTION

Feynman[1] was correct when he predicted "There's Plenty of Room at the Bottom." Since Feynman's talk, given at Caltech in 1959, there has been an explosion of research in the ever-shrinking areas of nanoscience. A nanometer (nano—Greek for *dwarf*) is one billionth of a meter, or ten times the size of an individual hydrogen atom. Since the mid-1980s, there has been a substantial increase in interest over the creation and use of nanocapsules,[2] a subset of nanoparticles. Nanocapsules have a number of potential applications, some of which include dye dispersants, nanoreaction vessels, and encapsulation media for fragrances, flavors, or drugs. This article will discuss the specific case of oil-filled nanocapsules. Such structures readily encapsulate lipophilic/hydrophobic and oil-soluble compounds, allowing them to be dispersed and transported through aqueous media.

DEFINING OIL-FILLED NANOCAPSULES

Nanoparticles have their largest dimension in the 1–200 nm range. Nanoparticles can be made of a wide range of materials: inorganic, organic, and biological. Nanoparticles are of interest in materials science for a number of reasons. Because chemical characteristics such as optical, magnetic, electrical, adsorptive, and catalytic are size-dependent, a nanoparticle may exhibit useful properties not seen in the larger bulk sample.[3] The surface area-to-volume ratio can be very high, which can increase the performance of catalysts.[3] For these reasons, material scientists are investigating a wide number of nanostructures, hoping to find new applications. One possible application is the encapsulation of guest moieties, which would be possible with hollow or filled nanoparticles. Nanocapsules are superior to nanospheres and other nanoparticles for these applications because of their high loading capacity.

The general definition of a nanocapsule is a spherical, hollow structure with a diameter less than 200 nm. The cavity can be empty or filled with a solvent, either polar or nonpolar (Fig. 1). Nanocapsules can be distinguished from other nanoparticles because they have a well-defined core and shell, whereas the latter do not. Although this definition includes coated nanoparticles, the term *nanocapsule* as referred to here is a submicrometer colloidal particle with a cavity surrounded by a shell. When made from polymers, nanocapsules have also been referred to as "hollow polymer nanostructures," "polymer/oil composite particles",[4] and "colloidosomes."[5,6] Nanocapsules have been made for many years, following the example of nature, using phospholipids, which are amphiphilic. Phospholipids will self-assemble in aqueous environments to form vesicles, which have a water core and a bilayer of the phospholipid forming the shell. This is the basis of cells. In a similar manner, a single layer of phospholipid can stabilize oil droplets, resulting in liposomes (Fig. 2).

Micro-Pak, Inc. has produced lipid vesicles that can be tailored to have either aqueous or oil-filled central cavities.[7] This central cavity can be filled wholly or in part with oil such that it can be used to encapsulate hydrophobic molecules.[7] The core of these vesicles is large and unstructured, which makes it ideal as a transport vehicle. Phospholipid vesicles have two major drawbacks.[7] First, phospholipids can be degraded by a number of enzymes, thus causing problems in biomedical applications. Second, the most common phospholipids contain polyunsaturated acyl chains, which are prone to peroxidation, resulting in a fracture of the vesicle. Liposomes are generally unstable.[8] Although similar to liposomes, polymeric capsules are held together through covalent bonding in the shell, making them more robust. Thus polymeric oil-filled nanocapsules are just robust, man-made mimics of naturally occurring liposomes.

The remainder of this review is divided into three major sections. The first deals with the characterization techniques used when making nanocapsules. The second deals with research published to date, and the third briefly describes some of the possible applications for oil-filled nanocapsules.

CHARACTERIZATION TECHNIQUES

Because of the small size and often complex formulations, a detailed characterization of nanocapsules is difficult. Some of the techniques used include quasi-elastic light scattering [QELS; also known as photon correlation spectroscopy (PCS)], gel permeation chromatography (GPC),

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013636

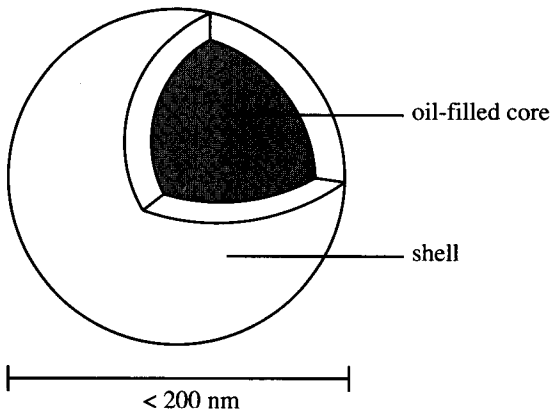


Fig. 1 Schematic of an oil-filled nanocapsule.

transmission electron microscopy (TEM), scanning force microscopy (SFM), and scanning electron microscopy (SEM). To date, none of these techniques has provided a clear picture of the nanocapsule shell.[9] It remains unknown as to whether the shell is a continuous or a porous network. The synthetic route used to attain the nanocapsules may also affect the nature of the shell. The characteristics of interest are morphology, size and size distribution, density, and zeta potential.

Morphology

Morphology refers to the structural shape of the nanocapsule. This can be determined using a number of microscopy techniques. The most commonly used is TEM, with either positive staining using a heavy atom such as osmium, or negative staining where the medium surrounding the nanocapsule is stained (typically with phosphotungstate or uranyl acetate). Another technique, which has proven effective, is freeze-fracture TEM, where the polymer shell, inner core, and wall thickness are all observed. Quintanar-Guerrero et al.[10] have already used SFM for oil-filled nanocapsules. Cryoelectron microscopy and SFM have a promising future for determining nanocapsule morphology and surface properties.[9] Scanning electron microscopy has also shown some applicability for examining the morphology of nanocapsules. It is limited by its magnification ability, but when combined with x-ray analysis, it is a powerful tool to identify nanocapsules with specific atoms. For example, it would be possible to characterize nanocapsules incorporating the iodine-containing oil, Lipiodol®.[9]

Size and Size Distribution

For many applications, nanocapsules of specific, uniform size are desirable. Microscopy techniques are capable of giving information about size and size distribution, but have some limitations. First, the typical sample size in microscopy is very small, thus one has to make the assumption that it is representative of the whole. Second, the sample is usually dried, which may adversely affect the morphology. To obtain data on the whole population, QELS or PCS is typically used. Nanoparticles dispersed in a fluid are in constant Brownian motion. The speed of the particles is inversely proportional to their size. Photon correlation spectroscopy analyzes the frequency of shifts in the light intensity that results from the particles moving into and out of the lightpath. Time dependence provides the information needed to determine the diffusion coefficient of the nanocapsules in solutions. Once the diffusion coefficient ($\langle D \rangle_z$) is determined, it can be related to the hydrodynamic radius (R_h) of the nanocapsule using the Stokes–Einstein law:

$$R_h = \frac{kT}{6\pi\eta\langle D\rangle_z} \quad (1)$$

where η is viscosity and kT is the Boltzmann coefficient multiplied by temperature. The distribution of the

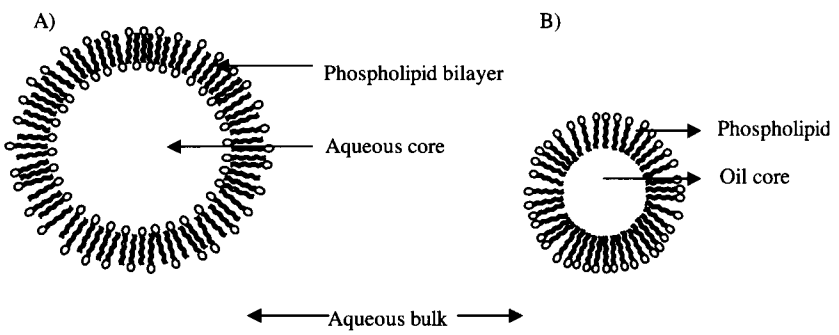


Fig. 2 Two-dimensional schematic of a 3-D phospholipid vesicle (A) and a liposome (B). The phospholipids shown actually form a shell around a 3-D sphere in each case.

frequency shifts in light intensity also provides the size distribution of the sample.

It is important to note that the size determined in this method is R_h, which includes a sphere of solvation around the nanocapsule. As a result, the size may appear larger than that observed in microscopy images.

Gel permeation chromatography can be performed to determine the molecular weight and polydispersity of the prepolymer.

Density

Chouinard et al.[11] showed that nanocapsule density could be determined by isopycnic centrifugation. This technique can be performed using either a classical or analytical centrifuge. Isopycnography (or equilibrium density gradient centrifugation) is a separation technique based on the differing intrinsic densities of the particles. The nanocapsules migrate through the centrifuge tube until they reach a point where they are surrounded by a solvent of equivalent density. The solvent used for nanocapsule analysis is Percoll™. The particle density can be backcalculated by comparing with a calibration curve obtained from density-calibrated particles, by collecting the zone of the gradient in which the particles accumulated, and by measuring its density.[11]

Comparing densities is a good method to determine whether the nanoparticles observed by microscopy are solid nanospheres, hollow nanocapsules, or oil-filled nanocapsules. The density of nanocapsules tends to be intermediate between that of the polymer matrix and the oil.[9] Solid nanospheres have a density equal to that of the polymer matrix. By comparing the density achieved via isopycnography with the densities of the starting materials, one can determine whether nanospheres or nanocapsules are present.[11] The density can also be used to yield the nanocapsule oil/polymer composition:[9]

$$1/\rho_{NC} = X_{poly}(1/\rho_{poly}) + X_{oil}(1/\rho_{oil}) \quad (2)$$

$$X_{poly} + X_{oil} = 1 \quad (3)$$

where X_{poly} and X_{oil} are the mass fractions of the polymer and oil in the nanocapsules, respectively, and represent unknowns. ρ_{NC}, ρ_{poly}, and ρ_{oil} are the densities of the nanocapsules, polymers, and oils, respectively.

Zeta Potential

Zeta potential is a surface characterization technique that can be used to investigate whether a guest molecule is truly encapsulated or simply adsorbed onto the surface of a nanocapsule.[9]

Particles, when dispersed in water, have a charge because of surface–solvent interactions. For example, silica particles lose a surface proton from silanol (Si–OH) groups to the aqueous solvent. The charged particle surface attracts a layer of counterions from the bulk solvent, leading to a double layer, whose thickness is dependent on the surface charge density. A large charge stops particles from getting close to each other because of electrostatic repulsion. Conversely, a small surface charge results in a smaller double layer, leading to flocculation.

Zeta potential is a measure (in millivolts) of the energy needed to shear the particle and its inner layer of counterions away from the bulk solvent. The higher the zeta potential is, the larger is the double layer and the more stable the nanocapsule is in the solution. Changes in zeta potential can indicate the adsorption of moieties to the surface of the nanoparticle, which have affected its double layer.

SYNTHETIC ROUTES

Two main techniques are used to synthesize nanocapsules: the first is interfacial polymerization around a droplet, and the second is deposition of a preformed polymer at the interface of a droplet. Both techniques utilize self-assembly.

Interfacial Polymerization

In this method, either a monomer or an amphiphilic polymer with a cross-linkable group is used and polymerization is induced at the surface of an oil droplet. This technique requires a system that can polymerize on a faster time scale than the fluctuation in the droplet geometry. The advantage of this technique is that the polymeric shell is formed in situ, allowing it to follow the contours of the droplet.[9] The drawback is that the cross-linking or polymerization reaction may have side reactions with other species present in the system. For example, if the nanocapsule is being used to encapsulate a drug, the monomer or cross-linkable group may have reactivity toward the drug. At the least, this would result in a reduced availability of the drug and, at its worst, this could result in toxic by-products.

Thurmond et al.[12] of the Washington University, St. Louis have prepared nanocapsules containing a hydrophilic shell and hydrophobic core properties. Thurmond et al. rely on self-assembly to form micelles, which are then stabilized through covalent bonding of the outer shell (Fig. 3). The result is a surface cross-linked "Knedel" (SCK).[12] The shell cross-linking provides two functions; first, it stabilizes the micelles; second, it offers the ability to control the permeability of the shell.[13] Thurmond et al.'s work was one of the first examples of chemically cross-linked nanosphere surfaces. Lowe et al.[14] and

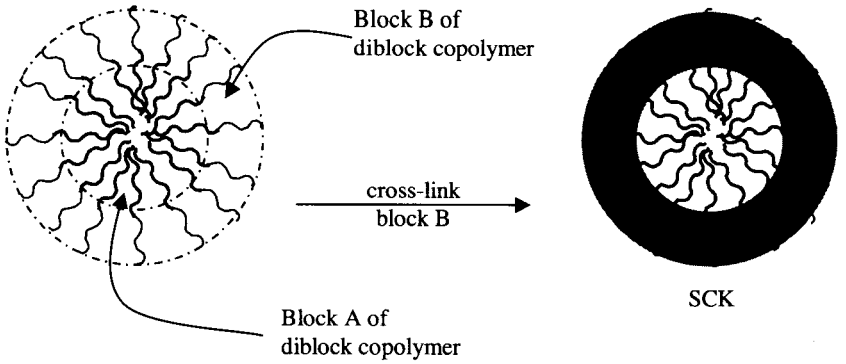


Fig. 3 Two-dimensional schematic of a 3-D SCK. (Adapted from Ref. [12].)

Bütün et al.[15,16] have also created SCK structures. In this case, the core could be reversibly hydrated or dehydrated depending on the solution temperature resulting in a hydrophilic or hydrophobic core, respectively.

Although SCKs have core–shell structures, they are not oil-filled. They are made from unswollen micelles, thus the core is composed of the hydrophobic, alkyl chain ends of the block copolymers used to make the shell.[17] Despite this, depending on the properties of the core polymer chains (glassy, fluidlike, or crystalline), the SCKs can have a variety of core properties, including the fluidlike properties found in oil-filled nanocapsules.[18] These last properties can be achieved by using a block copolymer that has a degradable core. Huan et al.[19] and Zhang et al.[20] showed the utility of this approach by hydrolyzing the poly-ε-caprolactone core of SCKs. The hydrolysis products resulted in a polar core, but it is conceivable that a similar process could be used with different polymers to yield nonpolar, oil-like cores.

The synthesis of oil-filled nanocapsules using amphiphilic surfactants follows a route similar to that used to make liposomes. Oil-filled nanocapsules are created when an oil-in-water microemulsion is used as the template for the interfacial polymer layer (Fig. 4). Micrometer-sized capsules have already been proven possible[21] from regular emulsions, and the chemistry can be extended to the formation of nanocapsules from microemulsions. Such nanocapsules are water-soluble, but coating the nanocapsules with additional layers of polymers. proteins, or other materials can change solubility. Surface materials are chosen for their specific properties, such as chemical and biological resistance, adhesion, and reactivity.

Water-in-oil emulsions have been used to make nanocapsules with aqueous cores.[22] The emulsions were stirred while a shell was polymerized at the oil–water interface. Similar techniques have been used to make nanocapsules with oil cores from oil-in-water microemulsions.[4,23–26] Underhill et al.[23] utilized a microemulsion stabilized with surfactants containing polymerizable head groups. Once the microemulsion has been formed, subsequent steps lead to the polymerization of the surfactant with a comonomer to form a robust shell around the oil droplet. McDonald et al.[24] also used microemulsions, but utilized the change in solubility between monomers and polymers to force a polymer shell to form at the interface between the oil droplet and the aqueous

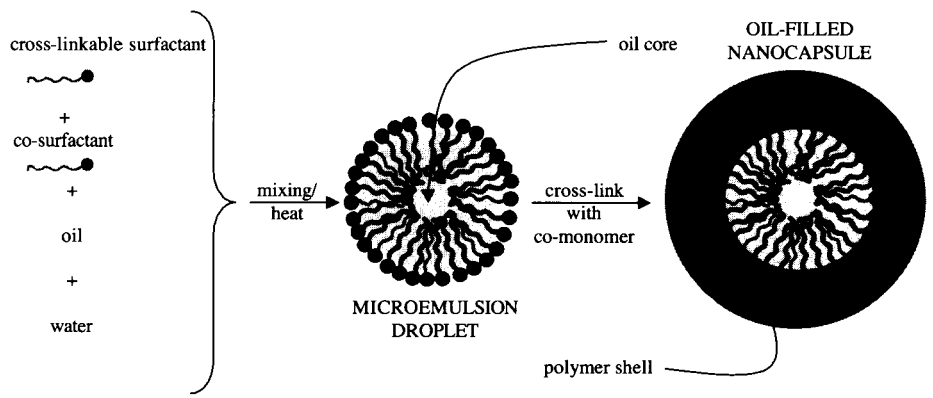


Fig. 4 Two-dimensional of the synthesis of nanocapsules using a microemulsion as a template. (From Ref. [23].)

Table 1 Research into nanocapsules via interfacial polymerization, classified by method, system and research group

System investigated	Research group	Refs.
SCK's-Polystyrene-block-poly(4-vinyl pyridine), Poly(2-(dimethylamino)ethyl methacrylate-block-methacrylic acid), Poly(ε-caprolactone)-block-poly(acrylic acid)	Wooley et al., Armes et al.	[12,14–20]
Monomer polymerization: Poly(alkylcyanoacrylate), Polystyrene, Poly(methacrylic acid), Poly(acrylonitrile)	Lenaerts et al., Couvreur et al., McDonald et al., Jang et al.	[11,22,24–26]
Surface cross-linkable surfactants: Polysiloxane/silicates	Duran et al.	[23]

continuous phase. The monomer was dissolved in the oil along with a hydrophobe and an initiator. The hydrophobe was chosen such that it was not miscible with the polymer; thus as the reaction progressed, the polymer was forced to segregate to the oil–water interface, leaving the hydrophobe forming the oil phase in the core. Typically, the hydrophobe is removed via vacuum or steam stripping. The hydrophobe diffuses through the shell to create hollow nanocapsules.[24] This is because the shell is porous enough to allow the hydrophobe to diffuse through the polymer. In the case of Jang and Ha,[26] the hydrophobe was isooctane, which is a lightweight volatile organic. If the hydrophobe is not removed, then oil-filled nanocapsules result (Table 1).

Preformed Polymer Assembly

The alternate method to form nanocapsules is to self-assemble preformed polymers around an oil droplet (Fig. 5). Again the droplet acts as a template. This method avoids the issue of the reactivity of the cross-linkable group and the possible side reactions. The drawback is that there is less control of the size and size distribution of the resulting nanocapsules.[9]

Nanocapsules, via the self-assembly of polymers at the interface of a droplet, can be made using natural (e.g., liposomes) and/or man-made [e.g., poly(ethyleneimine)] polymers. The deposition of a premade polymer at the surface of a droplet can be performed following a method introduced by Lvov et al.[27,28] in the mid-1990s. The shell wall of a nanocapsule is made via layer-by-layer assembly through alternate adsorption of oppositely charged components onto a template.[29] Uncharged polymers can be used if the precipitation method is changed. In this case, the polymer is dissolved in the organic phase, which is mixed under high shear with an aqueous phase containing a hydrophilic surfactant. As the nanometer-sized droplets of oil form, the polymer precipitates at the interface with the aqueous phase and the whole nanocapsule is stabilized by the surfactant. The resulting nanocapsules can be fine-tuned to provide the desired characteristics by changing the type and amount of polymers, oils, and/or surfactants. The ratio of the organic phase to the aqueous phase also plays a role in the chemical nature of the nanocapsule.[9]

The number of layers adsorbed to the surface determines the thickness of the shell. As an alternative to adsorbing polymers to the surface, colloidosomes have colloidal particles adsorbed to an emulsion droplet, resulting in a selectively permeable membrane, which allows submicron particles to diffuse in but excludes larger particles.[5,6,30–33] In this case, the resulting capsules exceed the size limit imposed on the definition of nanocapsules. Despite this limitation, it is plausible that this technique could be used for smaller entities. Of

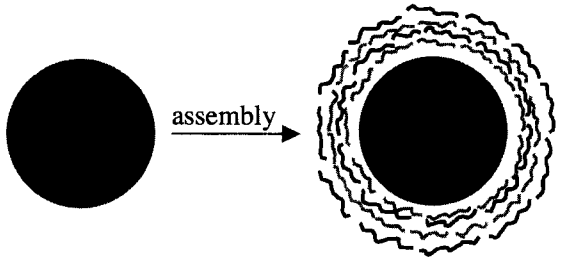


Fig. 5 Two-dimensional schematic of the layer-by-layer self-assembly of two different polymers (gray and black) on the surface of a template, which may be either a colloidal particle or an oil droplet. (Adapted from Refs. [27] and [28].)

Table 2 Research into nanocapsules via preformed polymer assembly, classified by method, system and research group

System investigated	Research group	Refs.
Emulsification-diffusion	Quintanar–Guerrero et al.	[10]
Colloidosomes	Dinsmore et al., Velev et al.	[5,6,30–33]
Silicate/polycation multilayers	Lvov et al.	[27,28]
Cationic/anionic polyelectrolyte multilayers	Decher	[29]

concern may be that with perfectly packed spheres, the pores are approximately $0.15d$, where d is the diameter of the adsorbed spheres. Thus as the spheres decrease in size, so will the basic pore size, resulting in decreased permeability (Table 2).

APPLICATIONS

Submicrometer-sized hollow particles have been explored for the encapsulation of a variety of guest molecules into their cores. They have an advantage over solid nanoparticles because their cores provide for high loading capacity. Different authors have referred to them differently; some examples include microspheres,[34,35] hollow nanoparticles,[36,37] silica particles,[38,39] nanocapsules,[11,40,41] solid lipid nanoparticles,[42] and host–guest carriers.[43,44] One of the advantages of oil-filled systems is that hydrophobic chemicals can be encapsulated and then transported at concentrations higher than their normal solubility in transport media. Supramolecular chemistry has involved numerous types of polymeric assemblies providing enhanced properties (e.g., stealth™, which is the ability of a nanoparticle to remain invisible to defense mechanisms within the body, and molecular recognition, which gives the nanoparticles the ability to target specific areas within the body). One of the limitations of such assemblies is their weak stability under changing environmental conditions. Therefore a combination of self-assembly followed by covalent bonding to create novel macromolecular architectures may be employed. The SCK nanostructured particles are being investigated for application in areas as broad as drug delivery, encapsulating agents from fragrances to food flavors, coatings, dye dispersants, pollutant removal systems, and catalysis. In the case of drug delivery, nanocapsules may be advantageous by providing protection against proteolytic degradation. In the case of the food and fragrance industries, nanocapsules are desirable because their use in encapsulating odors leads to a persistence of the odor long after application. This means a longer-lasting perfume, or a food product, which does not go "stale" as rapidly. Further variation over the three-dimensional (3-D) shape of polymer structures (e.g., the preparation of cylindrical or needle-shaped particles) is expected to generate materials of unique behavior.

As seen in the previous paragraph, oil-filled nanocapsules have a variety of applications. There are a number of articles in the literature, which outline their use in drug application technologies (see Refs. [9], [13], and [34] and the references therein). Some of the biologically active compounds that nanocapsules have been proposed to encapsulate are hormonal substances, antibiotics, insulin, proteins, antigens, viruses,[45] bacteria,[38] or cells.[46] One nonbiomedical application is in the cosmetic/dermatology industry. L'Oréal has been using liposomes since the early 1970s to encapsulate cosmetic ingredients.[47] Liposomes have several problems associated with their use in cosmetics.[48] First, encapsulation does not improve the bioavailability of the active ingredient. Second, liposome loading efficiency is low.[48] Last, the liposomes are not necessarily stable over all environments. Nanocapsules are superior to liposomes for this application because they address all these problems.

L'Oréal's nanocapsules have been on the market since 1995.[45,47–49] L'Oréal uses nanocapsules to encapsulate active ingredients such as vitamin A (retinol), vitamin E, vitamin C, and beta-carotene, which are most effective when they reach the deeper layers of the skin. Vitamin A is necessary for new cell generation and overall correct functioning of the epidermis.[48] Vitamin E "mops up" harmful oxygen free radicals. Both vitamins are more efficient when transported through the surface of the skin to the deep layers beneath. The nanocapsules facilitate transportation to these deep layers. The nanocapsules have a biodegradable shell, which is attacked by cutaneous enzymes, releasing the contents to the deep layers of the skin.[48]

The nanocapsules created by L'Oréal, in collaboration with the French National Research Ministry (CNRS), are made using a microemulsion technique. Two solvent phases are present, both being totally miscible with each other. The first solvent S_1 is a volatile liquid (e.g., acetone), which is used to dissolve the polymer (which will constitute the shell; e.g., poly-ε-caprolactone), the oil phase (which will be the core), the lipophilic active compound (which will be transported within the nanocapsule), and a surfactant (which will prevent the nanocapsules from agglomerating once they are formed). The second solvent S_2 is in aqueous phase and is a nonsolvent for the constituents dissolved in S_1. Both S_1 and S_2 are mixed together via a fine stream and nanometer droplets of oil are formed, surrounded by the polymer, which is insoluble in both S_1 and S_2, thus creating nanocapsules. The surface is coated with a stabilizing surfactant. As a final step, the acetone is removed under vacuum.[48]

A method of delivery alternate to the nanocapsules is the use of microemulsions. Nanocapsules are superior in terms of stability. Microemulsions can be destabilized by changes in the environment such as temperature and pH. Nanocapsules also provide increased bioavailability because their shells protect the active ingredient(s) in the core from cutaneous enzymes, which may degrade them. Nanocapsules have been shown to penetrate the epidermis further than microemulsions (up to 30 μm).[45] In the case of vitamin A, the nanocapsules act as reservoirs in the

uppermost layers of the epidermis, thus "stockpiling" the vitamin until the skin needs it.[48]

CONCLUSION

Oil-filled nanocapsules consist of spherical structures typically <200 nm in diameter with a distinct core–shell structure, where the core is fluidlike and lipophilic. Oil-filled nanocapsules can be synthesized by either the polymerization of a shell, or by the self-assembly of preformed polymers around a lipophilic compound. Differing monomers, polymers, and hydrophobes have resulted in various sizes and morphologies on nanocapsules. Oil-filled nanocapsules have been examined using SEM, TEM, and SFM.

Oil-filled nanocapsules are a rapidly developing technology that shows promise in a number of areas. The areas of application are broad, and include the areas of drug delivery[9,13] and removal,[23] food science,[50] cosmetics/dermatology,[47] paints/inks and coatings, pollutant removal systems, and vessels for catalysis (e.g., nanoreactors).[8] The commonality in all these applications is the need for a reservoir or host to encapsulate a guest moiety, which otherwise may not be soluble, or may be degraded by the environment in which it is needed. When used as nanoreactors, oil-filled nanocapsules can be considered as local areas of high concentration even if the overall solution concentration is low. This may help catalysis and reactivity.

Research groups currently working on nanocapsules for drug delivery acknowledge that the nanocapsules would be more effective if the surface were altered to allow for stealth abilities (stealth refers to the nanocapsules' ability to remain invisible to the body's defense mechanisms, e.g., white blood cells). This is possible through surface functionalization with poly(ethylene oxide). Nanocapsules would also be more efficient drug delivery vessels if they could target the specific site for drug delivery (e.g., a tumor). This is known as active targeting.[13] This can be achieved if the surface of the nanocapsule is modified with ligands, which are selectively recognized by receptors on the cells of interest (e.g., tumor cells). This is of importance because it provides maximum efficiency of the drug dosage and minimizes the side effects of the drug associated with it coming into contact with nontarget tissues.

In the future, nanocapsules will be better applied to topical, biomedical, cosmetic, and dermatological applications if a biomimetic approach is adopted. The nanocapsule structure will need to be engineered such that is resembles more closely the structure of single-celled organisms. The result would be a delivery device, which would deliver its contents to a specific activity site.

Nanocapsules have already shown their utility in a number of areas and continue to show promise with the possibility of adapting them to a number of circumstances with a little bit of imagination.

ACKNOWLEDGMENTS

The author would like to thank the members of the George and Josephine Butler Polymer Laboratory and the Engineering Research Center for Particle Science and Technology at the University of Florida as well as DRDC Atlantic for their support, both scientific and financial. Special thanks are given to Dr. R. S. Duran and Dr. G. Liu for their mentorship, and my colleagues at the DRDC Atlantic for their help and encouragement.

REFERENCES

1. Feynman, R.P. There's Plenty of Room at the Bottom. In *Engineering and Science*; Caltech, February 1960.
2. Al Kouri Fallouh, N.; Roblot-Treupel, L.; Fessi, H.; Devissaguet, J.P.; Puisieux, F. Development of a new process for the manufacture of polyisobutylcyanoacrylate nanocapsules. Int. J. Pharm. **1986**, *28*, 125–132.
3. Förster, S.; Antonietti, M. Amphiphilic block copolymers in structure-controlled nanomaterial hybrids. Adv. Mater. **1998**, *10*, 195.
4. Tiarks, F.; Landfester, K.; Antonietti, M. Preparation of polymeric nanocapsules by miniemulsion polymerization. Langmuir **2001**, *17*, 908–918.
5. Dinsmore, A.D.; Hsu, M.F.; Nikolaides, M.G.; Marquez, M.; Bausch, A.R.; Weitz, D.A. Colloidosomes: Selectively permeable capsules composed of colloidal particles. Science **2002**, *298*, 1006–1009.
6. http://www.deas.harvard.edu/projects/weitzlab/research/csome.html.
7. Wallach, D.F.H. Hybrid Paucilamellar Lipid Vesicles. US Patent 5,628,936, May 13, 1997.
8. Graff, A.; Winterhalter, M.; Meier, W. Nanoreactors from polymer-stabilized liposomes. Langmuir **2001**, *17*, 919–923.
9. Couvreur, P.; Barratt, G.; Fattal, E.; Legrand, P.; Vauthier, C. Nanocapsule technology: A review. Crit. Rev. Ther. Drug **2002**, *19* (2), 99–134.
10. Quintanar-Guerrero, D.; Allemann, E.; Doelker, E.; Fessi, H. Preparation and characterization of nanocapsules from preformed polymers by a new process based on emulsification–diffusion technique. Pharm. Res. **1998**, *15* (7), 1056–1062.
11. Chouinard, F.; Buczkowski, S.; Lenaerts, V. Poly

(alkylcyanoacrylate) nanocapsules: Physicochemical characterization and mechanism of formation. Pharm. Res. **1994**, *11*, 869–874.
12. Thurmond, K.B.; Kowalewski, T.; Wooley, K.L. Water-soluble Knedel-like structures: The preparation of shell-cross-linked small particles. J. Am. Chem. Soc. **1996**, *118*, 7239–7240.
13. Rösler, A.; Vandermeulen, G.W.M.; Klok, H.-A. Advanced drug delivery devices via self-assembly of amphiphilic block copolymers. Adv. Drug Deliv. Rev. **2001**, *53*, 95–108.
14. Lowe, A.B.; Billingham, N.C.; Armes, S.P. Synthesis and characterization of zwitterionic block copolymers. Macromolecules **1998**, *31*, 5991–5998.
15. Bütün, V.; Billingham, N.C.; Armes, S.P. Synthesis of shell cross-linked micelles with tunable hydrophilic/hydrophobic cores. J. Am. Chem. Soc. **1998**, *120*, 12135–12136.
16. Bütün, V.; Lowe, A.B.; Billingham, N.C.; Armes, S.P. Synthesis of zwitterionic shell cross-linked micelles. J. Am. Chem. Soc. **1999**, *121*, 4288–4289.
17. Thurmond, K.B.; Kowalewski, T.; Wooley, K.L. Shell cross-linked Knedels: A synthetic study of the factors affecting the dimensions and properties of amphiphilic core–shell nanospheres. J. Am. Chem. Soc. **1997**, *119*, 6656–6665.
18. Wooley, K.L. Shell cross-linked polymer assemblies: Nanoscale constructs inspired from biological systems. J. Polym. Sci., A, Polym. Chem. **2000**, *38*, 1397–1407.
19. Huan, H.; Remsen, E.E.; Kowalewski, T.; Wooley, K.L. Nanocages derived from shell cross-linked micelle templates. J. Am. Chem. Soc. **1999**, *121*, 3805–3806.
20. Zhang, Q.; Remsen, E.E.; Wooley, K.L. Shell cross-linked nanoparticles containing hydrolytically degradable, crystalline core domains. J. Am. Chem. Soc. **2000**, *122*, 3642–3651.
21. Dyatlov, V.A. Katz, G.A. Patent WO 94/17789, August 18, 1994.
22. Lambert, G.; Fattal, E.; Pintoalphandary, H.; Gulik, A.; Couvreur, P. Polyisobutylcyanoacrylate nanocapsules containing an aqueous core as a novel colloidal carrier for the delivery of oligonucleotides. Pharm. Res. **2000**, *17*, 707–714.
23. Underhill, R.S.; Jovanovic, A.V.; Carino, S.R.; Varshney, M.; Shah, D.O.; Dennis, D.M.; Morey, T.E.; Duran, R.S. Oil-filled silica nanocapsules for lipophilic drug uptake: Implications for drug detoxification therapy. Chem. Mater. **2002**, *14*, 4919–4925.
24. McDonald, C.J.; Bouck, K.J.; Chaput, A.B.; Stevens, C.J. Emulsion polymerization of voided particles by encapsulation of a nonsolvent. Macromolecules **2000**, *33*, 1593–1605.
25. Jang, J.; Lee, K. Microemulsion polymerization for hollow polymer nanosphere using encapsulation of a hydrophobe. Polym. Prepr. **2002**, *43*, 605–606.
26. Jang, J.; Ha, H. Fabrication of hollow polystyrene nanospheres in microemulsion polymerization using triblock copolymers. Langmuir **2002**, *18*, 5613–5618.
27. Lvov, Y.; Ariga, K.; Onda, M.; Ichinose, I.; Kunitake, T. Alternate assembly of ordered multilayers of SiO_2 and other nanoparticles and polyions. Langmuir **1997**, *13*, 6195–6203.
28. Lvov, Y.; McShane, M.; Jones, S.: de Villiers, M. In *Nanocapsule Technology Based on Layer-by-Layer Self-Assembly*, Abstracts of the Louisiana Conference on Commercial Applications of Microsystems, Materials and Nanotechnologies, Ruston, LA, USA, October 21–22, 2002.
29. Decher, G. Fuzzy nanoassemblies: Toward layered polymeric multicomposites. Science **1997**, *277*, 1232–1237.
30. Velev, O.D.; Lenhoff, A.M.; Kaler, E.W. A class of microstructured particles through colloidal crystallization. Science **2000**, *287*, 2240–2243.
31. Velev, O.D.; Furusawa, K.; Nagayama, K. Assembly of latex particles by using emulsion droplets as templates: 1. Microstructured hollow spheres. Langmuir **1996**, *12*, 2374–2384.
32. Velev, O.D.; Furusawa, K.; Nagayama, K. Assembly of latex particles by using emulsion droplets as templates: 2. Ball-like and composite aggregates. Langmuir **1996**, *12*, 2385–2391.
33. Velev, O.D.; Nagayama, K. Assembly of latex particles by using emulsion droplets: 3. Reverse (water in oil) system. Langmuir **1997**, *13*, 1856–1859.
34. Uhrich, K.E.; Cannizzaro, S.M.; Langer, R.S.; Shakesheff, K.M. Polymeric systems for controlled drug release. Chem. Rev. **1999**, *99*, 3181–3198.
35. Kramer, P.A. Albumin microspheres as vehicle for achieving specificity in drug delivery. J. Pharm. Sci. **1974**, *63*, 1646–1647.
36. Kreuter, J. Nanoparticle-based drug delivery systems. J. Control. Release. **1991**, *16*, 169–176. and references therein.
37. Brannon-Peppas, L. Recent advances on the use of biodegradable microparticles and nanoparticles in controlled drug delivery. Int. J. Pharm. **1995**, *116*, 1–9.
38. Finnie, K.S.; Bartlett, J.R.; Woolfrey, J.L. Encapsulation of sulfate-reducing bacteria in a silica host. J. Mater. Chem. **2000**, *10*, 1099–1101.

39. Bartlett, J.; Alexandr, B.C.J. AU Patent WO0162232, August 30, 2001.
40. Speiser, A. Controlled release of drugs from microcapsules and nanocapsules. Acta Pharm. Suec., Suppl. **1976**, *13*, 35.
41. Watnasirichaikul, S.; Davies, N.M.; Rades, T.; Tucker, I.G. Preparation of biodegradable insulin nanocapsules from biocompatible microemulsions. Pharmacol. Res. **2000**, *17*, 684.
42. Zimmerman, E.; Muller, R.H. Electrolyte- and pH-stabilities of aqueous solid lipid nanoparticle (SLN™) dispersions in artificial gastrointestinal media. Eur. J. Pharm. Biopharm. **2001**, *52* (2), 203–210.
43. Orr, G.W.; Barbour, L.J.; Atwood, J.L. Controlling molecular self-organization: Formation of nanometer-scale spheres and tubules. Science **1999**, *285*, 1049–1052.
44. Meissner, R.S.; Rebek, J.; deMendoza, J. Autoencapsulation through intermolecular forces: A synthetic self-assembling spherical complex. Science **1995**, *270*, 1485–1488.
45. Handjani, R.-M.; Ribier, A. Composition for the Cosmetic and/or Pharmaceutical Treatment of the Upper Layers of the Epidermis by Topical Application to the Skin, and Corresponding Preparation Process. US Patent 6,203,802, March 20, 2001.
46. Diaspro, A.; Silvano, D.; Krol, S.; Cavalleri, O.; Gliozzi, A. Single living cell encapsulation in nano-organized polyelectrolyte shells. Langmuir **2002**, *18*, 5047–5050.
47. Oger, G. The nanoscience behind beauty is serious business at L'Oréal. Smalltimes December 26, **2002**. (www.smalltimes.com).
48. Quin, F. In *Chemistry in Beauty*, Friday Evening Discourse at the Royal Institute of Great Britain, March 23, 2001. (www.rigb.org/events/transcripts.html).
49. Ribier, A.; Simonnet, J.-T.; Michelet, J. Cosmetic or Dermatological Composition Comprising an Oil-in-Water Emulsion Comprising Oily Globules with a Lamellar Liquid Crystal Coating. US Patent 6,066,328, May 23, 2000.
50. Gibbs, B.F.; Kermasha, S.; Alli, I.; Mulligan, C.N. Encapsulation in the food industry: A review. Int. J. Food Sci. Nutr. **1999**, *50*, 213–224.

Optical Molecular Devices

A. Prasanna de Silva
Nathan D. McClenaghan
School of Chemistry at Queen's University, Belfast, Northern Ireland

INTRODUCTION

Molecules (and supermolecules)[1,2] form the smaller size range of nano-objects, especially those that allow a degree of rational design including control of size or other properties and those that possess some useful function. This article puts the spotlight on molecules whose usefulness stems from the human comprehensibility of light signals. When a molecule is empowered with light absorption/emission, its small size can be an advantage to operate in tiny spaces but yet remain under a degree of human remote control. Most of these are sensing and switching devices,[3–5] the latter including some logic capabilities. Some of the design principles governing these will be outlined below. These are classified in terms of the formatting of chromophore/fluorophore and receptor components.[6] Chromophores and fluorophores are dyes that give light absorption signals and in the latter case, light emission signals as well. As their name suggests, receptors serve to receive species which are chemical in our cases. Thus chromo/fluorophores and receptors allow physical and chemical transactions, respectively.

"CHROMOPHORE–RECEPTOR" SYSTEMS

The most famous optical molecular devices of this kind are the pH indicators[7] known to every student of high school chemistry. Occupation of the receptor by a proton disturbs the electron distribution of the former. Because of the direct coupling between the chromophore and the receptor, it is therefore natural that the electron distribution of the chromophore itself is disturbed (Fig. 1). The consequence is a significant change of the absorption spectrum of the chromophore. Such acid-induced color changes have brightened up many a chemistry class around the world. A common example would be **1**. In general, such π-electron systems have electron donor and electron acceptor terminals, which leads to charge separation in the excited state. Such fractionally charged regions in these internal charge transfer (ICT) excited states can easily lead to observation of spectral wavelength shifts as a result of electrostatic interactions with the newly arrived target ion.

As simple as these indicators are, from a chemical standpoint, they also show a logic activity that has hitherto gone unnoticed. For instance, **2**[8] shows a simple blue shift of its absorption spectrum upon interaction with Ca^{2+}. However, this blue shift can be examined as a series of optical transmittance values obtained at different wavelengths of observation. As Fig. 2 shows, four wavelengths can be picked out to show clear Ca^{2+}-induced transmittance changes of the "low–high," "high–low," "low–low," and "high–high" variety. When "high" is coded as binary 1 and "low" is coded as binary 0, these digital input–output patterns can be identified as arising from single-input logic devices of the YES, NOT, PASS 0, and PASS 1 types, respectively. Furthermore, all of these logic behaviors can be simultaneously observed because light signals are readily multiplexed. So it is clear that humble ion indicators can show superposed logic behavior, which is unknown in the semiconductor device world.

"RECEPTOR$_1$–CHROMOPHORE–RECEPTOR$_2$" SYSTEMS

It is only logical to add another receptor to a "chromophore–receptor" system to develop more sophisticated formats (Fig. 3). This can be particularly productive when the two receptors are chosen to be selective, each to its own target species. For instance, **3**[8] takes in H^+ and Ca^{2+} at its quinoline nitrogen and amino acid receptors, respectively. The π-electron system again develops a dipole in the excited state with the positive pole being near the amino acid nitrogen and the negative end being close to the quinoline nitrogen. Thus admission of Ca^{2+} causes a destabilization of the excited state and hence a blue shift of the absorption spectrum (Fig. 4). On the other hand, the entry of H^+ causes a stabilization of the excited state and hence a red shift of the absorption spectrum. Of course, the simultaneous treatment of **2** with H^+ and Ca^{2+} gives a near cancellation of these spectral shifts. So an interesting

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009336

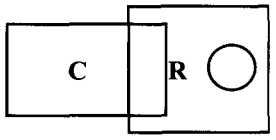
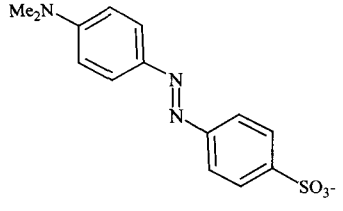


Fig. 1 The general format of a "Chromophore (C)–Receptor (R)" system.

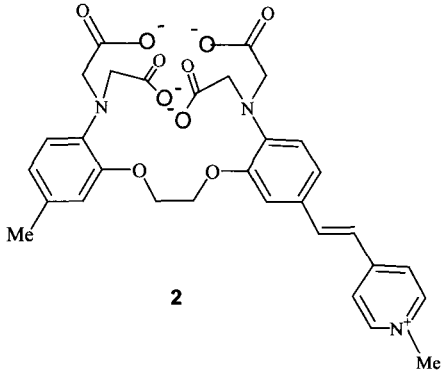
situation arises, where the spectral effect caused by two target ions is nearly the same as what is seen in their absence. Hence we can choose a monitoring wavelength where the transmittance of light is low (coded as binary 0) when the input target species are both low (H^+ and Ca^{2+} both coded as 0) or both high (H^+ and Ca^{2+} both coded as 1). Furthermore, each target ion on its own causes an absorption spectral shift away from the ion-free position. So now the transmittance of light is high (coded as binary 1) when the one input target species is low and the other high (H^+ coded as 1 and Ca^{2+} coded as 0 or its permutation). When these results are cast into a logic truth table (Fig. 3), we see that **3** behaves as a two-input XOR gate.

"FLUOROPHORE–RECEPTOR" SYSTEMS

Fluorescent versions of ion indicators[9] also have a long history and their mode of action borrows extensively from



1



2

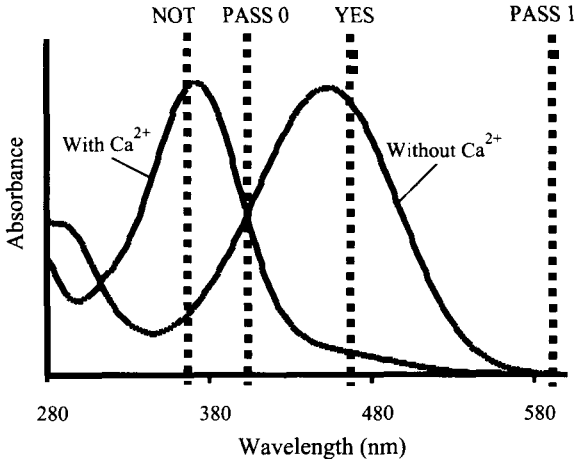
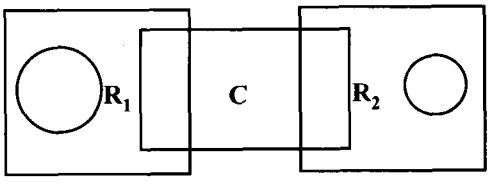


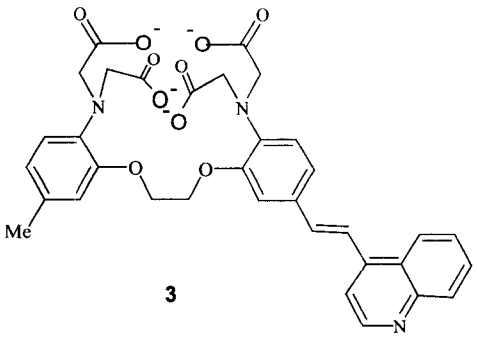
Fig. 2 Simultaneous observation of all four single-input logic types from a single experiment with a Ca^{2+} indicator.

their absorption-based cousins. One of the significant deviations of "fluorophore–receptor" systems arises as a result of the relative temporal delay before fluorescence emerges from an excited molecule (Fig. 5). Electrostatic repulsion between the photo-produced charge separations and the receptor-incumbent target species during this time period can cause decoordination of the target. Thus the target-induced spectral change will also dissolve away. Fluorescence emission spectra are therefore weakly influenced by target binding in many "fluorophore–receptor" systems known so far, although several exceptions are available. Of course, the target-induced changes survive in the fluorescence excitation spectra, which are related to the absorption spectra anyway. Grynkiewicz



$Input_1$ H^+	$Input_2$ Ca^{2+}	Output **Transmittance**
0	0	0
0	1	1
1	0	1
1	1	0

Fig. 3 The general format of a "$Receptor_1$ (R_1)–Chromophore (C)–$Receptor_2$ (R_2)" system and the logic truth table for the corresponding XOR gate.



3

et al.'s[10] excellent Ca^{2+} sensor **4** illustrates this very well. Here is an iconic optical molecular device that has served the cellular physiology community for nearly two decades now by imaging Ca^{2+} populations within living cells.

"FLUOROPHORE–SPACER–RECEPTOR" SYSTEMS

The apparently trivial addition of a spacer between a fluorophore and a receptor (Fig. 6) can completely change the device characteristics of the system. The spacer brings with it the ability to isolate components from the influence of short-range forces that normally abound in the chemical world. So the fluorophore and the receptor are forced to communicate via long-range interactions alone. These are few, and in many cases, can be reduced to one. Photoinduced electron transfer (PET), the celebrated mechanism of green plant photosynthesis, is the commonest controller of optical molecular devices of the "fluorophore–spacer–receptor" type. The fluorescence emission capability of the fluorophore is arrested by PET successfully competing for the energy of the excited state. Thus the device output is initially held in the "low" state (coded as 0). However, PET can be electrostatically stamped out, especially by charged target species when they take up residence in the receptor. Now excitation of the system will lead to no competition for the energy of the excited state. Consequently, the excited state returns to ground by emitting fluorescence as most fluorophores do. The device output is now "high" (coded as 1). Such target-induced fluorescence switching is logically a single-input YES gate. An example is the fluorescent sensor **5**[11] for Na^+, which is marketed by Roche Diagnostics for blood analysis in hospital critical care units. Cases such as **5** use more than electrostatics to enhance the fluorescence switching. The receptor within **5** is a *N*-(2-methoxyphenyl)monoaza-15-crown-5 ether, which suffers a major change in conformation upon capturing Na^+. This act reduces the electron delocalization within π-system of the receptor, which, in turn, makes the PET process more difficult and the fluorescence emission stronger. An extra feature within systems such as **5** is the ease with which components can be substituted for, in order to change the species being targeted or even its concentration range. So K^+-selective sensors with the same optical parameters found in **5** become available by simply replacing the receptor. Similarly, the Na^+-selective relatives of **5** such as **6**,[12] which communicate with different colors, of absorption and emission, are obtained by changing the fluorophore. Of course, the feasibility of PET must be conserved during such module replacements.

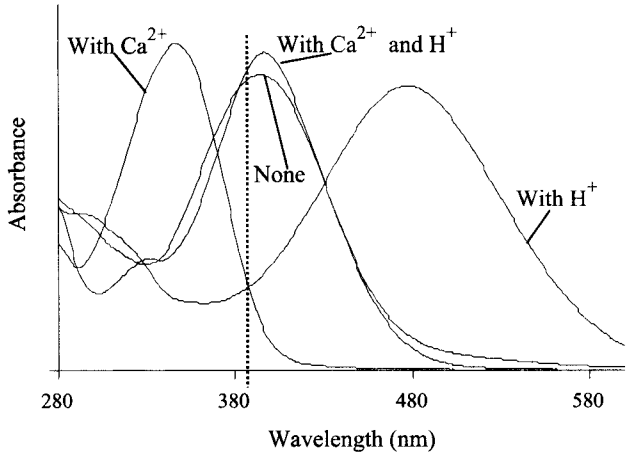


Fig. 4 The realization of general XOR logic behavior in the transmittance output at 390 nm of the UV–Vis absorption spectra set of **3**.

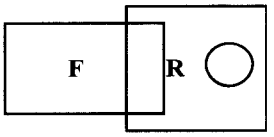


Fig. 5 The general format of a "Fluorophore (F)–Receptor (R)" system.

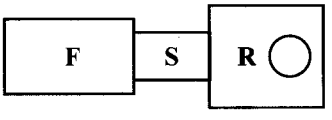
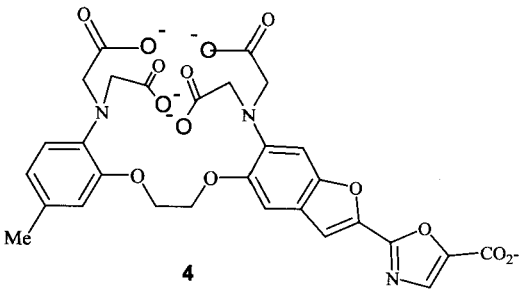
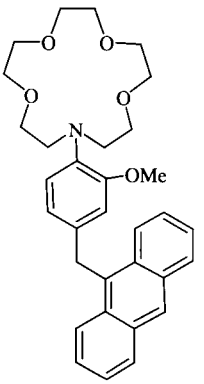


Fig. 6 The general format of a "Fluorophore (F)–Spacer (S)–Receptor (R)" system.



4

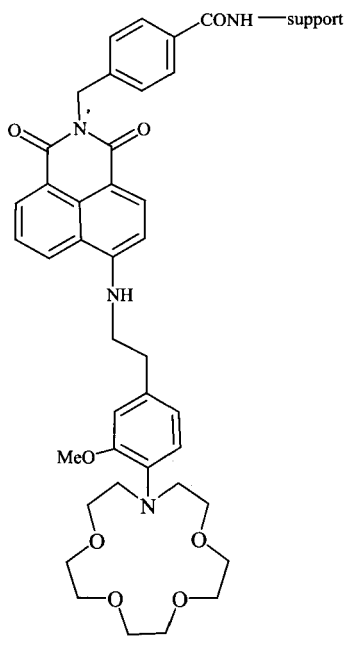


6

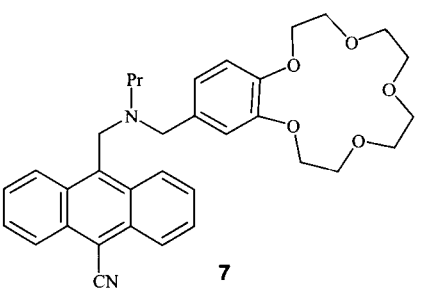
"FLUOROPHORE–SPACER$_1$–RECEPTOR$_1$–SPACER$_2$–RECEPTOR$_2$" SYSTEMS

As observed above, a spacer ensures a high degree of modularity of systems so that individual components are somewhat autonomous. This not only makes PET switch system design a predictive activity but also makes PET system expansion thoroughly logical. Addition of new modules will bring with them possibilities of PET, which are each predictable (provided that electron transfer data are available). Then it becomes possible to arrange situations in which two target species arrive at suitable receptors, either alone or together. Of course, we need to have adequate selectivity within the chosen receptors so that cross-talk of target species will be minimized. Now we have two-input, one-output devices that employ the same foundations as discussed above for one-input, one-output systems. In the simplest cases, "fluorophore–spacer$_1$–receptor$_1$–spacer$_2$–receptor$_2$" systems will have two possible PET paths originating from each receptor and finishing at the fluorophore, unless each is blocked by the correct target species. So fluorescence emerges unchallenged only if both receptors are blocked by the two target species being applied as inputs. The condition of Input$_1$=1 and Input$_2$=1 is required before a "fluorophore–spacer$_1$–receptor$_1$–spacer$_2$–receptor$_2$" PET system will pass an output=1. This is clearly AND logic (Fig. 7). The first example of this, and the first molecular logic gate of any kind in the primary literature, was **7**.[13] This uses H^+ and Na^+ as the two inputs. There are several excellent ways of arriving at molecular AND gates now,[14–18] some of which have led to more complex logical behavior.[14,19–21]

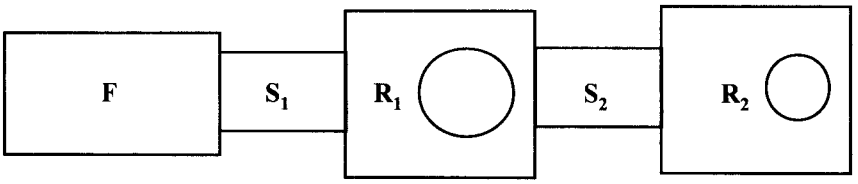
"Fluorophore–spacer$_1$–receptor$_1$–spacer$_2$–receptor$_2$" systems can also be put to uses that do not depend on binary logic. For instance, receptor$_1$ can be chosen as an electron donor amine, whereas receptor$_2$ can be chosen to be poorly electroactive. A pyridine is the choice for practical reasons, where the fluorophore is an anthracene unit within **8**.[22] Naturally, amines lose their electron



5

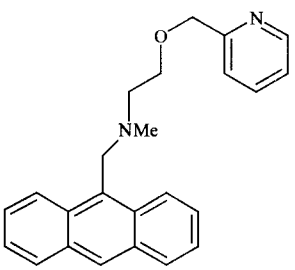


7



$Input_1$	$Input_2$	Output **Fluorescence**
0	0	0
0	1	0
1	0	0
1	1	1

Fig. 7 The general format of a "Fluorophore (F)–$Spacer_1$ (S_1)–$Receptor_1$ (R_1)–$Spacer_2$ (S_2)–$Receptor_2$ (R_2)" system and the logic truth table for the corresponding AND gate.



8

donor activity upon binding to a proton. PET processes are suppressed. On the other hand, pyridines become good electron acceptors upon proton binding. PET processes are created. So the proton target species has opposite effects upon arrival at the two receptors, each with its own concentration threshold for reception. We note that a single-input species causes a single fluorescence output to be controlled in a relatively complex way. At low proton concentrations, both receptors are free and the amine launches a PET process to destroy fluorescence. At mid-range proton concentrations, the more avid amine receptor picks up a proton, thereby closing its PET channel. Fluorescence flares up as a consequence. At high proton concentrations, both receptors are protonated. The protonated amine remains PET-disabled, but the newly formed pyridinium launches its own PET channel and extinguishes the fluorescence. Thus the fluorescence output follows a "off-on-off" pattern in response to monotonically ramping proton concentrations (Fig. 8).[22–24] Such systems are useful in being direct optical indicators of pH conditions of enzyme activity or even cellular activity. After all, the principle of "the happy medium" or "the middle way" affects everyone.

"$Receptor_1$–$Spacer_1$–Fluorophore–$Spacer_2$–$Receptor_2$" Systems

Realizable permutations arise when a sufficiently large number of modules are contained in a system. This is the case with "$fluorophore–spacer_1–receptor_1–spacer_2–receptor_2$" systems. A realizable permutation is to shift the fluorophore to the center of the system (Fig. 9). This act has an advantage for chemical design, because fluorescence switching efficiencies can be improved via accelerated PET processes arising from the shorter fluorophore–receptor distances involved. An example is **9**.[25] Even "$receptor_1–spacer_1–fluorophore–spacer_2–re-$

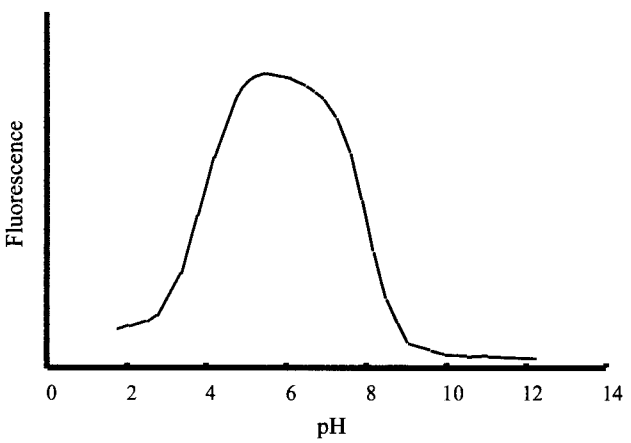


Fig. 8 The 'off-on-off' fluorescence–pH profile of **8**.

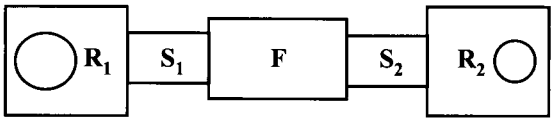
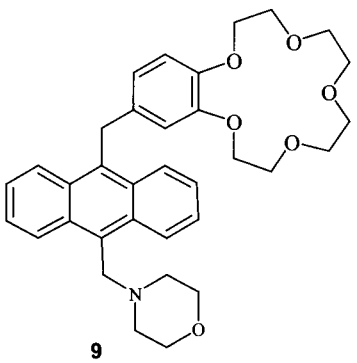


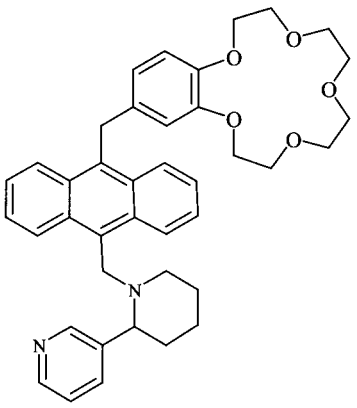
Fig. 9 The general format of a "Receptor$_1$ (R_1)–Spacer$_1$ (S_1)–Fluorophore (F)–Spacer$_2$ (S_2)–Receptor$_2$ (R_2)" system.

ceptor$_2$–spacer$_3$–receptor$_3$" systems are now in the hands of designers[26] to perform increasingly complex tasks with deceptively small molecules such as **10**.

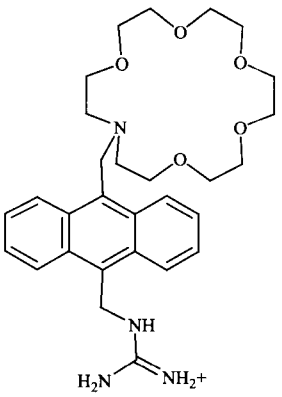
The cases discussed above involved separate target species such as Na^+ and H^+ arriving essentially simultaneously at their respective receptors. Of course, these target species can be independently controlled to test all the input combinations for setting up truth tables to assign logic behavior. It is also feasible to build, say, two target species into separate sites of a bifunctional molecule. Now the real target becomes the bifunctional molecule itself. Naturally, such bifunctional reception can lead to enhanced selectivity of binding and detection. For example,



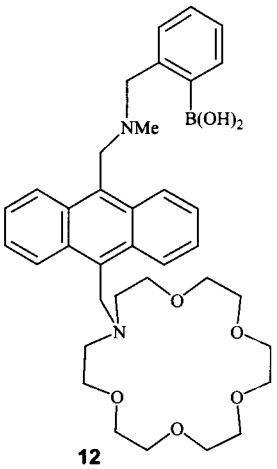
9



10



11



12

11[27,28] selectively targets amino acid zwitterions with a specified number of carbon atoms in between the ammonium and carboxylate functionalities. While the binding is enhanced, the fluorescence signaling suffers from a weakness. Indeed, the binding of the ammonium group leads to PET suppression and fluorescence enhancement. However, the capture of the carboxylate moiety reaps no such fluorescence reward owing to the lack of sufficient PET activity in the guanidinium group. Nevertheless, a nice case with two PET-active receptors is available in the form of **12** from Cooper and James.[29] Protonated glucosamine is the valuable target. The aza-18-crown-6 ether receives an ammonium group as in the case of **11**. Additionally, a diol feature is held by the aminomethylboronic acid receptor, which leads to PET suppression. So now both PET channels are blocked upon arrival of the glucosamine species in an AND logical

manner. This application of AND logic systems for the enhanced binding and optical signaling is a very promising avenue of research.

CONCLUSION

Dyes (fluorescent or not), receptors, and spacers are the building blocks that designers of optical molecular devices can play with. Combinations, or even some permutations, of these blocks can lead us to sensors, logic gates, and "off-on-off" systems already. Considering that at least some of these are demonstrably useful here and now, the number of players is bound to increase. The result will be even more interesting systems in the future.

ACKNOWLEDGMENTS

We thank the Department of Employment and Learning, Northern Ireland and the European Union (HPRN-CT-2000-00029) for support of our efforts in this area.

REFERENCES

1. Lehn, J.-M. *Supramolecular Chemistry*; VCH: Weinheim, 1995.
2. Balzani, V.; Scandola, F. *Supramolecular Photochemistry*; Ellis-Horwood: Chichester, 1991.
3. de Silva, A.P.; Gunaratne, H.Q.N.; Gunnlaugsson, T.; Huxley, A.J.M.; McCoy, C.P.; Rademacher, J.T.; Rice, T.E. Signaling recognition events with fluorescent sensors and switches. Chem. Rev. **1997**, *97*, 1515–1566.
4. *Chemosensors of Ion and Molecule Recognition*; Czarnik, A.W., Desvergne, J.-P., Eds.; Kluwer: Dordrecht, 1997.
5. Valeur, B. *Molecular Fluorescence*; Wiley-VCH: Weinheim, 2001.
6. Bissell, R.A.; de Silva, A.P.; Gunaratne, H.Q.N.; Lynch, P.L.N.; Maguire, G.E.M.; Sandanayake, K.R.A.S. Molecular fluorescent signalling with 'fluor–spacer–receptor' systems—Approaches to sensing and switches devices via supramolecular photophysics. Chem. Soc. Rev. **1992**, *21*, 187–195.
7. *Indicators*; Bishop, E., Ed.; Pergamon Press: Oxford, 1972.
8. de Silva, A.P.; McClenaghan, N.D. Simultaneously multiply-configurable or superposed molecular logic systems composed of ICT (internal charge transfer) chromophores and fluorophores integrated with one- or two-ion receptors. Chem. Eur. J. **2002**, *8*, 4935–4945.
9. Kirkbright, G. Fluorescent Indicators. In *Indicators*; Bishop, E., Ed.; Pergamon Press: Oxford, 1972; 685–708.
10. Grynkiewicz, G.; Poenie, M.; Tsien, R.Y. A new generation of Ca^{2+} indicators with greatly improved fluorescence properties. J. Biol. Chem. **1985**, *260*, 3440–3450.
11. He, H.; Mortellaro, M.A.; Leiner, M.J.P.; Young, S.T.; Fraatz, R.J.; Tusa, J.K. A fluorescent chemosensor for sodium based on photoinduced electron transfer. Anal. Chem. **2003**, *75*, 549–555.
12. Gunnlaugsson, T.; Nieuwenhuyzen, M.; Richard, L.; Thoss, V. Novel sodium-selective fluorescent PET and optically based chemosensors: Towards Na^+ determination in serum. J. Chem. Soc., Perkin Trans. 2 **2002**, 141–150.
13. de Silva, A.P.; Gunaratne, H.Q.N.; McCoy, C.P. A molecular photoionic and gate based on fluorescent signalling. Nature **1993**, *364*, 42–44.
14. Remacle, F.; Speiser, S.; Levine, R.D. Intermolecular and intramolecular logic gates. J. Phys. Chem., B **2001**, *105*, 5589–5591.
15. Lukas, A.S.; Bushard, P.J.; Wasielewski, M.R. Ultrafast molecular logic gate based on optical switching between two long-lived radical ion pair states. J. Am. Chem. Soc. **2001**, *123*, 2440–2441.
16. Stojanovic, M.N.; Mitchell, T.E.; Stefanovic, D. Deoxyribozyme-based logic gates. J. Am. Chem. Soc. **2002**, *124*, 3555–3561.
17. Pina, F.; Roque, A.; Melo, M.J.; Maestri, I.; Belladelli, L.; Balzani, V. Multistate/multifunctional molecular-level systems: Light and pH switching between the various forms of a synthetic flavylium salt. Chem. Eur. J. **1998**, *4*, 1184–1191.
18. Inouye, M.; Akamatsu, K.; Nakazumi, H. New crown spirobenzopyrans as light- and ion-responsive dual-mode signal transducers. J. Am. Chem. Soc. **1997**, *119*, 9160–9165.
19. de Silva, A.P.; McClenaghan, N.D. Proof-of-principle of molecular-scale arithmetic. J. Am. Chem. Soc. **2000**, *122*, 3965–3966.
20. Raymo, F.M. Digital processing and communication with molecular switches. Adv. Mater. **2002**, *14*, 401–409.
21. Raymo, F.M.; Giordani, S. Multichannel digital transmission in an optical network of communicating molecules. J. Am. Chem. Soc. **2002**, *124*, 2004–2007.
22. de Silva, A.P.; Gunaratne, H.Q.N.; McCoy,

C.P. Direct visual indication of pH windows: 'Off-on-off' fluorescent PET (photoinduced electron transfer) sensors switches. Chem. Commun. **1996**, 2399–2400.
23. de Silva, S.A.; Zavaleta, A.; Baron, D.E.; Allam, O.; Isidor, E.V.; Kashimura, N.; Percarpio, J.M. A fluorescent photoinduced electron transfer sensor for cations with an off-on-off proton switch. Tetrahedron Lett. **1997**, *38*, 2237–2240.
24. Fabbrizzi, L.; Licchelli, M.; Poggi, A.; Taglietti, A. A versatile fluorescent system for sensing of H^+, transition metals, and aromatic carboxylates. Eur. J. Inorg. Chem. **1999**, 35–39.
25. de Silva, A.P.; Gunaratne, H.Q.N.; McCoy, C.P. Molecular photoionic AND logic gates with bright fluorescence and ''off–on'' digital action. J. Am. Chem. Soc. **1997**, *119*, 7891–7892.
26. de Silva, S.A.; Amorelli, B.; Isidor, D.C.; Loo, K.C.; Crooker, K.E.; Pena, Y.E. A fluorescent 'off-on-off' proton switch with an overriding 'enable–disable' sodium ion switch. Chem. Commun. **2002**, 1360–1361.
27. de Silva, A.P.; Gunaratne, H.Q.N.; McVeigh, C.; Maguire, G.E.M.; Maxwell, P.R.S.; O'Hanlon, E. Fluorescent signalling of the brain neurotransmitter gamma-aminobutyric acid and related amino acid zwitterions. Chem. Commun. **1996**, 2191–2192.
28. Sasaki, S.; Hashizume, A.; Citterio, D.; Fujii, E.; Suzuki, K. Fluororeceptor for zwitterionic form amino acids in aqueous methanol solution. Tetrahedron Lett. **2002**, *43*, 7243–7245.
29. Cooper, C.R.; James, T.D. Synthesis and evaluation of D-glucosamine-selective fluorescent sensors. J. Chem. Soc., Perkin Trans. 2 **2000**. 963–969.

Optical Nanosensors and Nanobiosensors

Brian M. Cullum
University of Maryland, Baltimore County, Baltimore, Maryland, U.S.A.

INTRODUCTION

Advances in nanotechnology and nanofabrication have begun to have profound effects on many different scientific fields in recent years. In the field of chemical and biological sensing, one of the most significant advances in the last several decades has been the development of nanosensors, for the probing of microscopic environments. Simply stated, nanosensors are devices with dimensions on the nanometer scale that are capable of monitoring the presence of a specific chemical or class of chemicals. Although many different types of nanosensors (i.e., optical, electrochemical, etc.)[1–6] have been reported in the literature over the last two decades, this chapter will focus on nanosensors employing optical transduction methods.

As with larger optical sensors, optical nanosensors can generally be classified into one of two different classes: 1) chemical nanosensors, or 2) nanobiosensors, depending on the type of recognition element (i.e., chemical or biochemical) used to provide specificity to the sensor.[7–10] Although both of these classes of optical nanosensors are capable of obtaining quantitative measurements in many different microscopic environments, they have found an ideal application in the analysis of chemical and biochemical species present within living cells. Their small sizes allow them to be inserted and precisely positioned within individual cells to obtain spatially localized measurements of chemical species in real time.

OVERVIEW

Prior to the development of optical nanosensors, chemical analyses inside individual living cells were limited almost entirely to the field of fluorescence microscopy, wherein a fluorescent indicator dye is introduced into a cell and allowed to diffuse throughout. When the dye comes in contact with the analyte of interest, a change in the fluorescence properties (e.g., intensity, spectral shift, etc.) of the dye occurs and fluorescence images of the entire cell are obtained. From these images, it is possible to monitor the presence of the analyte at various locations. However, because this technique relies on imaging of the fluorescent dye, its homogenous dispersion throughout the cell is required. Unfortunately, homogeneous dispersion of dyes is often limited by intracellular conditions (i.e., pH, etc.), or prevented entirely by compartmentalization of the dye by the cell.[11] Therefore nanosensors offer significant improvements over such analyses in many cases, as they allow the user to obtain measurements at whatever location is desired without the need for homogenous dispersion of a fluorescent indicator dye.

Because of their potential for providing an enhanced understanding of cellular responses to various stimuli, several reviews have already been devoted to the subject of optical nanosensors, despite their short existence.[5,12–16] This chapter will look at the evolution of optical nanosensors from their beginning (near-field optical microscopy) to the present (biosensors capable of probing subcellular compartments of individual cells) and discuss their application to biological measurements, as well as future directions in optical nanosensing.

EVOLUTION OF OPTICAL NANOSENSORS

Presently, the most widely reported class of optical nanosensors is fiber optic nanosensors. These sensors employ fiber optics that have been tapered on one end to diameters typically ranging between 20 and 100 nm. Although such sensors are based on the same principles as larger, conventional fiber optic sensors, their excitation process is quite different. Because the diameter of the tapered end of a fiber optic nanosensor is significantly smaller than the wavelength of light used for excitation, photons cannot escape from the tip of the fiber to be absorbed by the species of interest, as is the case in larger fiber optic sensors. Instead, excitons or evanescent fields continue to travel through the remainder of the tapered fiber's tip, providing the necessary excitation energy. Because of the weak nature of these phenomena, excitation using such a sensor is highly localized, allowing only species close to the fiber's tip to be excited. The use of tapered fiber optics to produce highly localized excitation processes (i.e., near-field excitation) arose from an area of research known as near-field optical scanning microscopy (NSOM).

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009344

Near-Field Scanning Optical Microscopy

Near-field optical scanning microscopy is a relatively recent field of research, employing light sources and/or detectors that are smaller than the wavelength of light used for imaging.[17–20] By using such small excitation sources or detectors, images of a sample can be obtained with nanometer-scale (i.e., 10–50 nm) spatial resolution. One common method for performing such experiments is to place a pinhole in front of the detector, effectively reducing the detector's size to nanometer dimensions.[20,21] However, a variation of this technique that is growing more and more popular is to construct an excitation probe with dimensions that are smaller than the wavelength of light that is being used for sample interrogation, thereby providing a light source with subwavelength dimensions.[22,23] The first such excitation probe was developed by Betzig et al.[22] by tapering a single-mode optical fiber to dimensions of approximately 20 nm, thereby confining the excitation radiation to the fiber tip. Using this nanometer-scale tapered optical fiber, NSOM images of a known pattern were reconstructed with signal enhancements of greater than 10^4 [22,23] over previous NSOM analyses.[18,19,24,25]

Fiber tapering methods

To construct tapered fiber optic probes for NSOM, two different processes have been developed: heated pulling and chemical etching, each capable of creating fibers with various tip diameters, taper angles, and smoothness. The first and most commonly employed of these procedures uses a heated pulling instrument, such as a laser-based micropipette puller. In this procedure, a fiber optic is placed in the micropipette puller, and a CO_2 laser heats the fiber while it is stretched along its major axis by a tension device (Fig. 1). As the fiber is pulled, the heated region

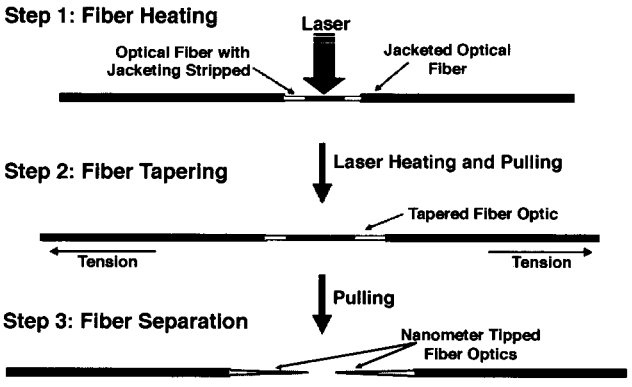


Fig. 1 Cartoon depicting the heated pulling process used for nanofiber fabrication. (*View this art in color at www.dekker.com.*)

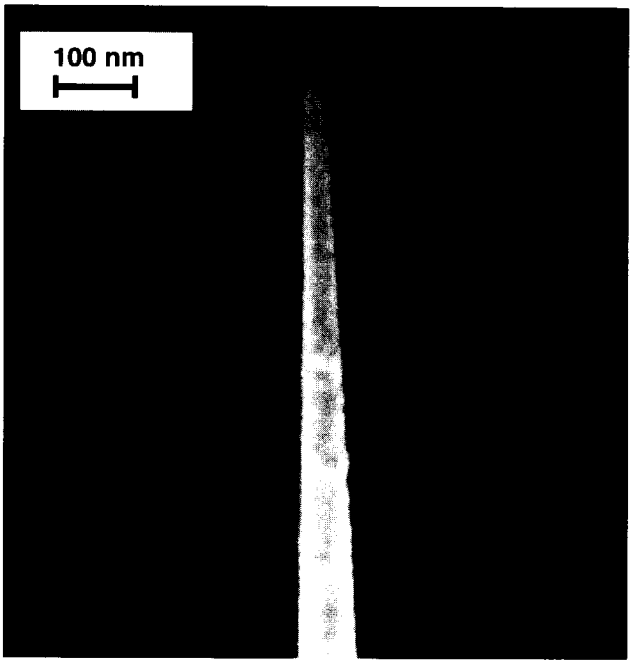


Fig. 2 A SEM of a tapered fiber optic probe. The diameter of the tapered tip is approximately 40 nm.

begins to taper until, finally, the fiber is pulled into two pieces, each having one large end and one end with nanometer-scale dimensions. By varying the heating temperature as well as the tension applied during the pulling process, tip diameters ranging in size from less than 20 nm to greater than 1000 nm have been reported.[22,26–29] A scanning electron micrograph (SEM) of an optical fiber that has been pulled with such an instrument is shown in Fig. 2. Using this technique, it is possible to produce optical fibers with highly reproducible nanometer-scale tips in just seconds.

In addition to the heated pulling process, optical fibers with nanometer-scale tips can also be produced via chemical etching. Two different variations of chemical etching have been reported in the literature: 1) Turner etching,[30,31] and 2) tube etching.[32] In both processes, hydrofluoric acid (HF) is used to etch the silica core of a fiber optic to a point. In the Turner method, the fiber is placed in the meniscus between HF and an organic overlayer, causing the HF to etch the silica and the organic overlayer to protect it. This process creates fibers with larger taper angles than those produced via the heated pulling method. These larger taper angles allow excitation light to travel closer to the tip of the fiber before being trapped, providing a more efficient excitation process. Unfortunately, because of the dual chemical nature of the etchant solution, environmental parameters such as temperature fluctuations and vibrations can cause significant batch-to-batch variations in the characteristics of the fiber tips.

Unlike Turner etching, the second chemical etching process, tube etching, employs a single-phase solution, thereby reducing the potential batch-to-batch variability induced by two phases. In this process, a silica core fiber having a polymer cladding is polished optically flat and one end is placed into a solution of HF. After being inserted into the HF, the silica core of the fiber begins to dissolve without affecting the cladding. Over time, the silica core continues to be etched away until it is no longer below the surface of the HF solution. When this occurs, capillary action draws HF up the walls of the polymer cladding. This HF then reaches the silica core of the fiber, where it drains back into the solution below, forming a smooth tip with a large taper angle (Fig. 3) within approximately 2 hr. To remove the surrounding polymer cladding and to expose the silica tip, the fiber is then simply placed in a suitable organic solvent. By varying the diameter of the fiber used and the depth at which it is inserted into the etching solution, accurate control over the fiber tip diameter and its taper angle can be achieved. Although this process can produce more reproducible tapered fiber optic probes than Turner etching, difficulties associated with submerging multiple fibers to the exact same depth in the HF solution make it less reproducible than the heated pulling process.

Application of fiber optic nanoprobes to near-field optical microscopy

One of the most significant applications of fiber optic nanoprobes to NSOM analyses of biological samples occurred when a single dye-labeled DNA molecule was detected using near-field surface-enhanced resonance Raman spectroscopy (NFSERRS).[33,34] In that work, dye-labeled DNA strands were spotted onto a surface-enhanced Raman spectroscopy (SERS) substrate that was prepared by evaporating silver on a nanoparticle-coated surface.[33,34] Following preparation of the sample, a fiber optic nanoprobe was raster-scanned over the sample's surface, illuminating it point by point, while the resulting Raman signals were measured with a charge-coupled device (CCD). Based on the intensity of the Raman signals measured at every location, a two-dimensional image of the DNA molecules was reconstructed and normalized for surface topography based on the intensity of the Rayleigh scatter.

With the advent of fiber optic nanoprobes, NSOM promises to be an area of research that holds a great deal of promise for biological analyses. Because many biological compounds produce luminescent signals following excitation, which are typically much stronger than Raman signals, it may be possible to map out the location of individual molecules of specific chemicals in human tissues (e.g., neurotransmitters in the brain) using NSOM. Such analyses could open new horizons in the investigation of complex chemical reactions and pathways of biological systems.

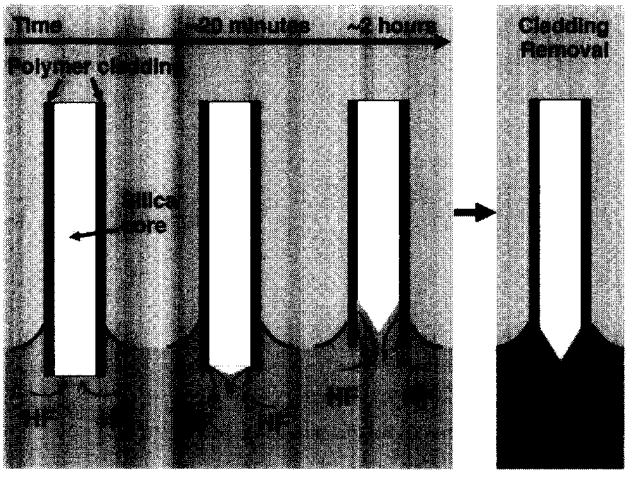


Fig. 3 Cartoon depicting the tube etching process used for nanofiber fabrication. (*View this art in color at www.dekker.com.*)

FIBER OPTIC CHEMICAL NANOSENSORS

Shortly after the development of fiber optic nanoprobes for NSOM, they were applied to the field of chemical sensing, making it possible to obtain spatially localized measurements (i.e., less than 50 nm away from the fiber's tip) of specific chemicals in three-dimensional structures.[26] Unlike simple NSOM probes, these fiber optic chemical nanosensors have chemical recognition elements (e.g., fluorescent indicator dyes, etc.) bound to the tapered tip of the fiber to provide a degree of specificity. Because of this specificity and their ability to obtain spatially localized measurements, the monitoring of concentration gradients and spatial inhomogeneities in submicroscopic environments (e.g., cells, etc.) with chemical nanosensors is possible. However, because of the small sampling volume probed by such fiber optic chemical nanosensors, the amount of analytes present in the excitation volume at any given time is very small, making it important to use a sensitive spectroscopic analysis technique (e.g., fluorescence, etc.).

In addition to employing a sensitive spectroscopic technique, it is also important to employ a sensitive detection system, such as the one shown in Fig. 4. In such a system, the sample is excited by launching an intense light source (e.g., laser) into the proximal end of the fiber optic nanosensor. The nanosensor is then positioned in the desired location using an x–y–z micromanipulator or

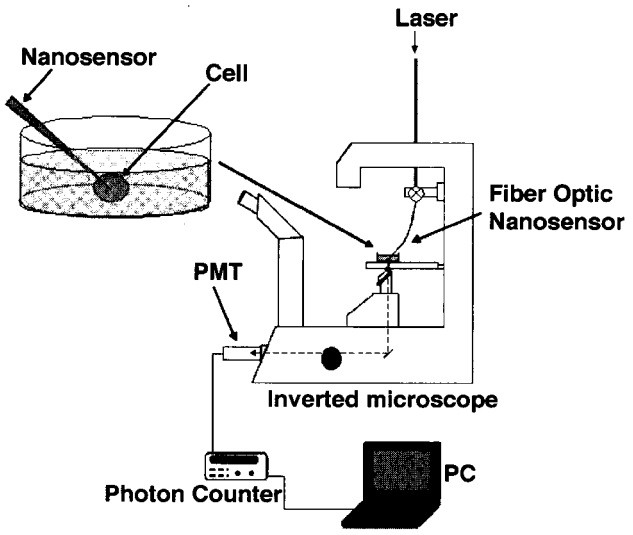


Fig. 4 Schematic diagram depicting a typical measurement system for fiber optic nanosensor-based analyses. The nanosensor is manipulated to the location of interest using an *x–y–z* micromanipulator. (*View this art in color at www.dekker.com.*)

piezoelectric positioning system mounted on a microscope. Once in place, the fluorescent indicator dye immobilized on the tip of the fiber is excited, and the resulting fluorescence emission is collected and filtered by the microscope before being detected with either a photomultiplier tube (PMT) or a CCD.

Ion-Sensitive Fiber Optic Nanosensors

The first fiber optic-based chemical nanosensors were reported in 1992 for the monitoring of pH.[26] In that original work, multimode fibers and single-mode fibers were tapered via a heated pulling process to produce tip diameters ranging from 100 to 1000 nm.[26,27] These tapered fibers were then used to fabricate pH-sensitive nanosensors via a three-step process. Immediately following the tapering process, a thick layer of aluminum was applied to the walls of the fibers using a vacuum evaporation system. This aluminum layer ensured that total internal reflection occurred over the tapered region of the fibers, where the cladding was stretched too thin. To prevent aluminum from depositing on the tips of the fibers, where the receptor molecules (i.e., fluorescent indicator dye molecules) were to be bound, the fibers were placed in an evaporator system with their tapered tips facing away from the source of the evaporating aluminum. This allowed the sides of the fibers to shadow the tips from the evaporating metal. This is illustrated in Fig. 5. Once the walls of the fibers were coated with aluminum, the silica tips were silanized to allow cross-linking to a polymer coating. The third and final step in the fabrication of these pH nanosensors consisted of attaching a pH-sensitive dye, acrylofluoresceinamine. to the silanized fiber tips through a variation of a photo-polymerization process that has been used in the construction of larger chemical sensors.[35,36] However, unlike the larger non-tapered fiber optic chemical sensors, the near-field excitation provided by these small fiber probes restricted the cross-linking of the polymer solution to the near-field of the fiber, providing a small sensing volume.

Following fabrication, the response characteristics of the pH nanosensors were evaluated. During this evaluation, the sampling volume of the nanosensors was determined to be more than six orders of magnitude smaller than conventional fiber optic chemical sensors, making them ideal for biological analyses. In addition, these nanosensors were found to be both stable and reversible to pH changes, and have response times of approximately 300 msec, which is approximately 100-fold faster than conventional fiber optic chemical sensors.

The application of fiber optic chemical nanosensors to a biologically relevant analysis was first performed by monitoring the pH of the extraembryonic space of rat conceptuses in an attempt to understand the role environmental factors play on embryonic growth.[26] In this study, pH nanosensors similar to the ones described above were inserted into the extraembryonic space of rat conceptuses, with minimal damage to the surrounding visceral yolk sac, and pH measurements were obtained. In a similar

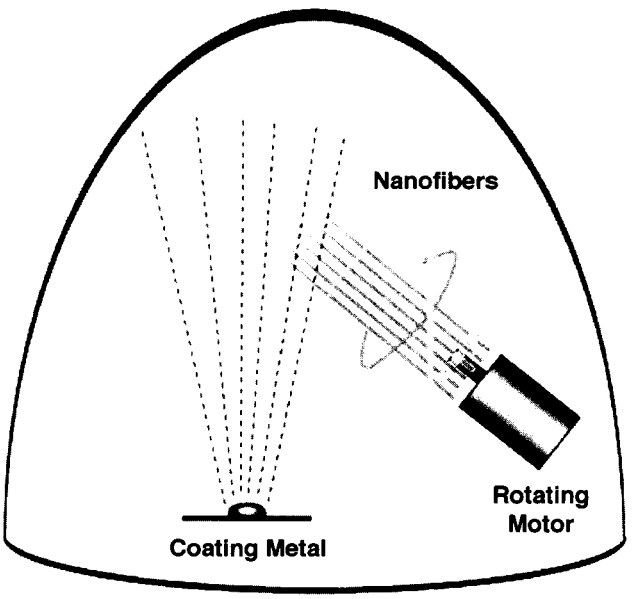


Fig. 5 Cartoon depicting the thermal evaporation procedure used to coat the sides of the tapered fiber optic probes with metal.

study using the same pH-sensitive nanosensors, indirect analyses of nitrite and chloride levels in the yolk sac of rat conceptuses were also performed.[37]

Since the construction of the first fiber optic nanosensors, additional fiber optic-based chemical nanosensors have been reported for the measurement of pH,[38–41] various metal ions,[42,43] as well as nitric oxide.[44] In addition, the size of the environments in which they probe continues to become smaller. In fact, the measurement of chemical species inside individual living cells has even been reported using fiber optic chemical nanosensors. The first such intracellular measurements with fiber optic chemical nanosensors employed sodium ion (Na^+)-sensitive indicator dyes immobilized onto the tips of tapered fiber optics through a process similar to that described for the pH nanosensors. These sensors were then used to measure sodium ion (Na^+) concentrations in the cytoplasmic space of a single mouse oocyte, one of the largest mammalian cells (ca. 100 μm in diameter), while ion channels were opened and closed by external stimulation.[15] In addition to intracellular analyses of Na^+ ions, fiber optic chemical nanosensors have also been used to monitor intracellular calcium (Ca^{2+}) ion fluctuations in vascular smooth muscle cells, following stimulation.[45] Because of the large number of cellular reactions that are signaled by fluctuations in the concentration of ionic species, and the demonstrated ability of fiber optic nanosensors to perform intracellular measurements of those ionic species, these sensors should prove to be valuable tools for future discoveries in cellular dynamics.

FIBER OPTIC NANOBIOSENSORS

Because of the complexity of biological systems and the number of possible interference to chemical nanosensors, the need for added specificity in cellular analyses can arise. To achieve this added specificity, fiber optic nanobiosensors are often employed. Like their larger counterparts, conventional fiber optic biosensors, biological receptor molecules (i.e., antibodies, enzymes, etc.) are used to provide added specificity. The different types of bioreceptor molecules that have been used for the fabrication of fiber optic nanobiosensors include antibodies, oligonucleotides, and enzymes, thereby allowing for the detection of a wide array of analytes.

Antibody-Based Fiber Optic Nanobiosensors

The first fiber optic nanobiosensor was reported by Alarie and VoDinh[28] in 1996. In this work, antibody-based nanobiosensors for the DNA adduct, benzo[*a*]pyrene tetrol (BPT), were developed and characterized. The fabrication of these first nanobiosensors was performed in a process very similar to that used in the development of fiber optic chemical sensors. Initially, a 600-μm-diameter multimode fiber optic was tapered down to 40 nm at the tip. After tapering, a thick layer of silver was applied to the sides of the fiber in such a way as to prevent the coating of the fiber's tip. Next, the uncoated fiber tip was silanized, and antibodies were attached via a covalent-binding procedure.[28] Following attachment of the antibodies, the fiber optic nanobiosensors were characterized in terms of antibody binding affinity as well as sensitivity and absolute detection limits. From these measurements, it was found that the antibodies had retained greater than 95% of their native binding affinity for BPT after being bound and that the absolute detection limit for BPT using these nanobiosensors was approximately 300 zmol (i.e., 300×10^{-21} mol).

Shortly after the development of this first antibody-based fiber optic nanobiosensor, several others were developed and applied to in vitro measurements within individual living cells.[5,29,46–48] In one such study, nanobiosensors for BPT were prepared as described above and used to obtain quantitative measurements of intracellular concentrations of BPT in the cytosol of two different cell lines: 1) rat liver epithelial cells, and 2) human mammary carcinoma cells.[29] Unlike previous intracellular analyses that employed relatively large cells (i.e., mouse oocytes and neurons), the cells analyzed in this study were spherical in shape and had diameters of approximately 10 μm, thereby demonstrating that fiber optic nanosensors and nanobiosensors could be used to analyze cells the size of typical mammalian somatic cells (i.e., 10–15 μm) without destroying them. In fact, this study also demonstrated that the insertion of a fiber optic nanosensor or fiber optic nanobiosensor into such a cell and the subsequent measurement seemed to have little effect on the cell's normal function. This was demonstrated by inserting the fiber optic nanobiosensor into the cytoplasm of a cell beginning to undergo mitosis and monitoring cell division following incubation of the nanofiber in the cell for 5 min. Fig. 6 contains an image of the fiber optic nanosensor being inserted into this cell, as well as a second image showing the two daughter cells that it divided into after approximately 2 hr.

Molecular Beacon-Based Fiber Optic Nanobiosensors

To detect the presence of oligonucleotides (i.e., RNA, DNA, etc.) in intracellular environments, a class of fiber optic nanobiosensors that employ a relatively new type of bioreceptor molecules known as a molecular beacon has been developed.[49] Molecular beacons are hairpin-shaped oligonucleotide probes that rely on the complementarity of nucleic acids (i.e., adenine:thymine, cytosine:guanosine, etc.) to form the basis for the molecular recognition

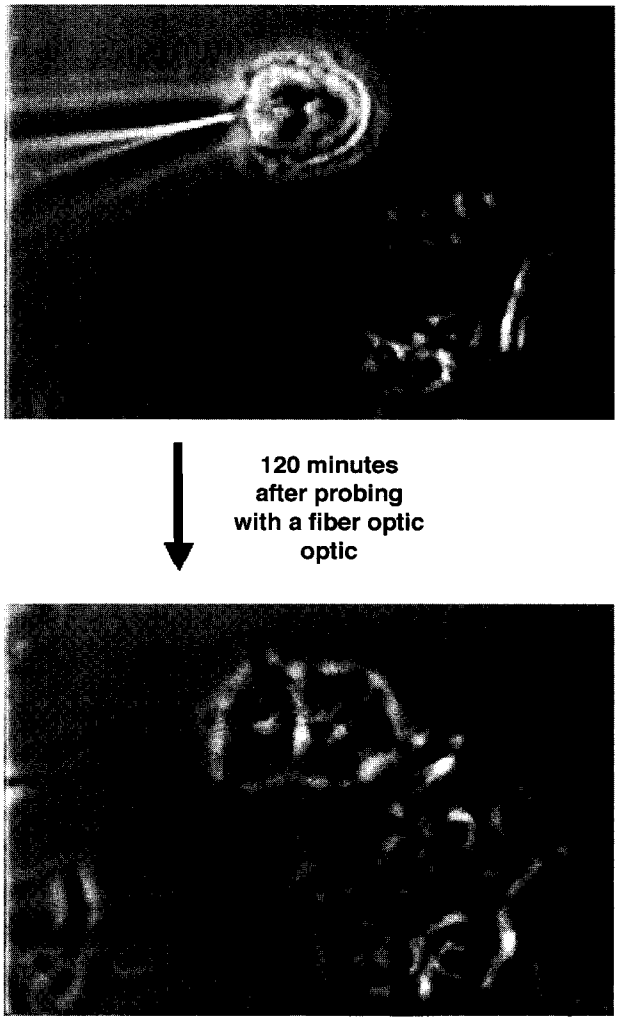


Fig. 6 Photograph of an individual cell that is probed with a fiber optic nanosensor as it begins to undergo mitosis, and a second photograph of that same cell approximately 2 hr later, having divided into two daughter cells. (*View this art in color at www.dekker.com.*)

of a specific oligonucleotide sequence. As shown in Fig. 7, molecular beacons form a stem-loop structure in the absence of the target sequence, in which a fluorophore [e.g., fluorescein isothiocyanate (FITC), etc.] on one end of the stem is in close proximity to a nonfluorescing, quenching moiety [e.g., dimethylaminophenylazobenzoic acid (DABCYL), etc.] on the other end.[50] In this state, when the fluorophore on the molecular beacon is excited, an energy transfer takes place between the excited fluorophore and the quencher, either by direct energy transfer or fluorescence resonance energy transfer (FRET), and minimal fluorescence emission occurs. However, when the loop sequence of the molecular beacon comes into contact with its complementary sequence, the stem begins to unzip, creating a spatial separation between the fluorophore and the quencher. This spatial separation in turn causes a dramatic increase in the fluorescence intensity of the molecular beacon, with some reports stating enhancements exceeding 200-fold.[51]

Because of their inherent sensitivity and specificity, molecular beacons have recently been used as bioreceptor molecules in the fabrication of fiber optic nanobiosensors for single-stranded DNA and RNA.[52,53] In these works, Liu and Tan and Liu et al. attached molecular beacons to the silica core of a fiber optic nanosensor through biotin–avidin linkages. By binding the molecular beacons to the tapered tip of the fiber optic in this method, a strong attachment was achieved, while not degrading the activity of the beacon.

Enzyme-Based Fiber Optic Nanobiosensors

In addition to the use of molecular beacons and antibodies as bioreceptor molecules, fiber optic nanobiosensors that employ enzymes for molecular recognition have also been developed.[54,55] Using enzymes as bioreceptors not only provides nanobiosensors with a high degree of specificity, but their catalytic activity can amplify the species being measured, allowing for sensitive analyses. One such enzymatic-based fiber optic nanobiosensor was developed for the indirect detection of glucose.[55] In this work, Rosenzweig and Kopelman immobilized the enzyme, glucose oxidase, and the oxygen-sensitive indicator, tris(1,10-phenanthroline) ruthenium chloride, in an acrylamide polymer on the tapered end of a nanofiber via a photo-polymerization process. Therefore when the nanobiosensor is in the presence of glucose, the enzyme catalyzes the oxidation of glucose into gluconic acid, consuming oxygen. The resulting changes in oxygen levels

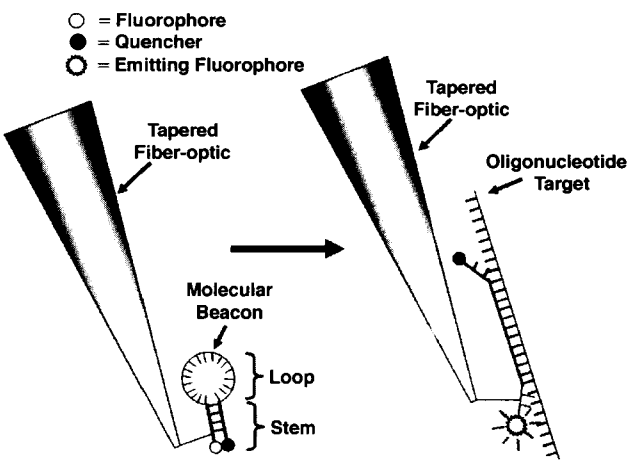


Fig. 7 Cartoon depicting the mechanism of action of a molecular beacon. (*View this art in color at www.dekker.com.*)

are then measured via the oxygen-sensitive indicator dye. By using an enzymatic receptor, these nanobiosensors were capable of absolute detection limits of approximately 10^{-15} mol and a sensitivity five to six orders of magnitude greater than current glucose optodes.[55]

Another example of an enzymatic nanobiosensor was developed by Tan et al. for the indirect measurement of the neurotransmitter, glutamate. In this work, glutamate dehydrogenase is bound to the tip of the fiber optic nanobiosensor to achieve molecular recognition. When the sensor is in the presence of glutamate, the nonfluorescent species, NAD, is reduced into the autofluorescent species, NADH, which can then be monitored via fluorescence spectroscopy. By relating the intensity of the NADH fluorescence to glutamate concentration, sensitive analyses were performed. Following the construction of these nanobiosensors, their usefulness was evaluated by continuously monitoring the release of glutamate from individual neurons during stimulation.

FIBER OPTIC NANOIMAGING SENSORS

Although fiber optic nanosensors and nanobiosensors have proven to be useful tools for the spatially localized, quantitative analysis of chemical species in cells, only one sensor can be inserted into a typical mammalian somatic cell at a time, without causing significant damage.[56] This inability to monitor multiple locations simultaneously can dramatically limit the applicability of fiber optic nanosensors and nanobiosensors to many types of cellular analyses. To overcome this limitation, a couple of research groups have recently begun to develop nanoscale fiber optic imaging probes and sensors comprised of thousands of individual fibers arranged in a coherent bundle.[57,58] These nanoimaging sensors represent the latest advance in fiber optic-based nanosensors, and are capable of obtaining measurements at thousands of different locations simultaneously with nanometer-scale spatial resolution.

Fiber Optic Nanoimaging Probe Fabrication

Similar to single-fiber-based nanosensors and nanobiosensors, nanoimaging probes can be fabricated either through a chemical etching process, or a heated pulling process. In the chemical etching process, commercially available coherent imaging bundles with 30,000 individual fibers are etched in HF to form an array of silicon tips.[58] Following formation of the array of silicon tips, a layer of gold is sputtered over the entire surface to prevent light from escaping from the tapered sides of each fiber element. After deposition of the gold overlayer, a polymer is electrochemically deposited over the gold and heat-cured to expose the gold at the tips of the individual fiber elements. Finally, the exposed gold surface on each of the individual fiber element tips is dissolved to produce free silica surfaces for spectroscopic probing.

In addition to chemical etching, nanoimaging probes have also been fabricated using a heated pulling process.[57] However, unlike the nanoimaging array created via chemical etching, the heated pulling process produces an array having a flat surface. This can be seen in Fig. 8A, which shows a side-on view of a fiber optic imaging bundle that has been tapered in a micropipette puller. To produce this flat surface, which is necessary for imaging, the delay between the final pull in the micropipette puller and the time at which the heating laser turned off is increased, thus allowing the fiber to cool off slightly and be broken, with all of the individual fiber elements being of the same length and diameter. By changing the heating temperature and the initial pull strength of the micropipette puller, fibers can be accurately and reproducibly tapered to all different diameters. Fig. 8 shows a fiber that has been tapered to have individual fiber element diameters of 800 nm on one end (Fig. 8B) and 4-μm diameters on the untapered end (Fig. 8C).

Using these two different fabrication techniques, fiber optic nanoimaging probes and sensors that will not only

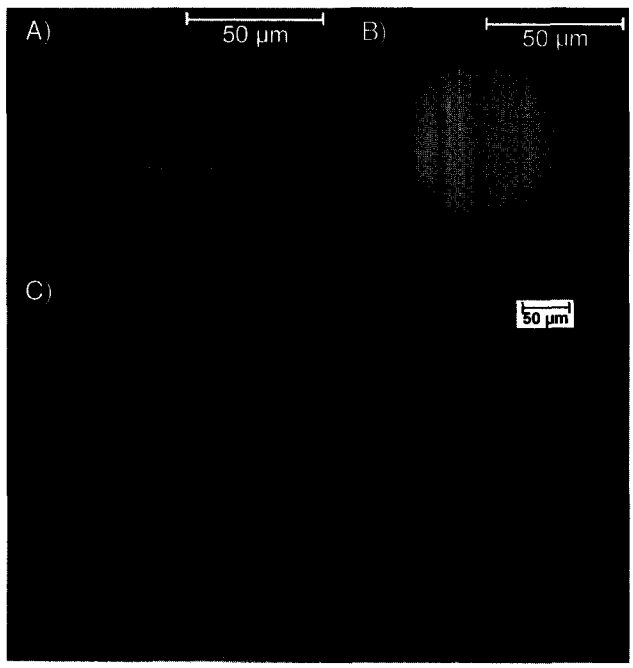


Fig. 8 Photographs of a fiber optic nanoimaging probe that has been tapered using a heated pulling process. Panel A shows a side-on view of the fiber demonstrating the flat tip after pulling. Panels B and C are end-on views of the tapered end and the nontapered end of the fiber, respectively. (*View this art in color at www.dekker.com.*)

allow for the quantitative measurement of individual chemical species, but also the simultaneous measurement at many different locations are being developed. Although these fiber optic imaging bundles are too large to be inserted into individual living cells, they can be used to monitor cellular membranes without damage.

NANOPARTICLE-BASED OPTICAL NANOSENSORS

Although fiber optic-based nanosensors have had a large impact on the fields of cellular biology and biochemistry, significant advances and variations in optical nanosensors are constantly being made, from the use of more selective bioreceptors to the development of different types of optical-based nanobiosensors. One such advance in the last several years has been the development of nanoparticle-based optochemical sensors, with nanometer-scale sizes in all three dimensions. Because of the small sizes of these sensors, a large number of them can be implanted within an individual cell at one time, allowing for the monitoring of many locations simultaneously. Although many different nanoparticle-based sensors are currently being developed, three main classes have already shown a great deal of promise for intracellular analyses. These three classes are quantum dot-based nanobiosensors, polymer-encapsulated nanosensors known as PEBBLEs, and phospholipid-based nanosensors. Each of these different classes of nanoparticle-based sensors is described in more detail below.

Quantum Dot-Based Nanosensors

The first class of nanoparticle-based optical nanosensor that was developed involved attaching nanometer-scale semiconductor particles, known as quantum dots, to various biological receptor molecules including antibodies, oligonucleotides, and enzymes.[59–63] The quantum dots used in these nanosensors comprised ZnS particles capped with Cd–Se. By using quantum dots instead of conventional fluorescent dyes, these sensors exhibit much more intense emission as well as a much greater degree of photostability. These properties are very important when trying to monitor changes in chemical or biochemical species concentrations over time because most fluorescent dyes exhibit rapid and significant photobleaching with the small amount used in cellular analyses. Unfortunately, because quantum dot-based nanobiosensors exhibit luminescence emission whether bound to the analyte or not, there are limits to their applicability to cellular analyses. Additionally, a biocompatibility issue also exists for these sensors because the materials used to fabricate the quantum dots (e.g., Cd–Se, etc.) are toxic to cells. Because of their toxicity, there is a significant effort currently underway in many research groups to develop more biocompatible quantum dot-based nanobiosensors to perform long-term monitoring of cellular reactions or processes.

PEBBLEs Sensors

A second class of nanoparticle-based optical nanosensors that has already had a large impact on the fields of cellular biology and biochemistry was developed by Sasaki et al.,[64] Clark et al.,[65,66,68] Sumner et al.,[67] and Xu et al.[69,70] in the last several years. These sensors, known as PEBBLEs (probes encapsulated by biologically localized embedding), are comprised of fluorescent indicator dyes sensitive to ionic species (i.e., H^+, Ca^{2+}, etc.) that are embedded in 20-nm-diameter or 200-nm-diameter polymer or sol–gel spheres.[65,66] By encapsulating the fluorescent indicator dyes in a polymer matrix, they are protected from cellular degradation by proteins, while still allowing ions to pass and react with them. In addition, this polymer coating also protects the cell from the toxic effects of the dye.

Since they were first developed in 1996, PEBBLEs have been applied to the measurement of many different species (pH, Ca^{2+}, NO, O_2, and Zn^{2+}) within individual cells.[65–70] In these analyses, large quantities of PEBBLEs are inserted in a cell, either by a gene gun or a similar device, to ensure that sensors are present at every location in which a measurement might be desired. Once the sensors have been injected into the cell, the entire cell is illuminated and the fluorescence signal from the indicator dye is measured over the autofluorescent background of the cell.

Because of the potential for intense autofluorescence from the cell, depending on the excitation wavelength used, the detection limits of such analyses can be relatively high. To overcome this problem, Anker et al.[71] and Anker and Kopelman[72] have recently developed magnetically modulated variations of these sensors, known as MagMOONs (magnetically modulated optical nanoprobes). These MagMOONs are created around aspheric magnetic nanoparticles that can be rotated in the presence of a rotating magnetic field. As the particles begin to rotate, the optical emission from the sensor is modulated. During demodulation, the emission from the sensors can be separated from the continuous autofluorescence background signals, thereby dramatically improving the detection limit of such sensors.

Phospholipid Sensors

The third major class of nanoparticle-based sensors that has already demonstrated a significant impact on cellular

analyses is phospholipid-based sensors. Like the PEBBLEs described above, these nanosensors also employ an encapsulation technique to ensure biocompatibility with the cell being investigated.[73–78] Within this class of nanosensors, two distinct subclasses exist: liposome sensors,[77,78] and lipobead sensors.[73,75] The first of these subclasses, liposome-based nanosensors, employs fluorescent indicator dyes encapsulated in the internal aqueous compartment of a liposome. This allows the dye, which is sensitive to a particular analyte, to retain its solution-based characteristics (i.e., spectral emission profile, Stokes shift, response time, etc.) while preventing toxic dye molecules from diffusing throughout the cell. Currently, liposome nanosensors have been developed and applied to the measurement of molecular oxygen[78] and pH[77] in various cellular environments, demonstrating a high degree of sensitivity as well as specificity.

Recently, a variation of these liposome-encapsulated nanosensors was developed by Ji et al.,[73] McNamara et al.,[75] McNamara and Rosenzweig,[79] and DeCoster et al.[80] in which the fluorescent indicator dye molecules were immobilized onto a polystyrene nanoparticle prior to being encapsulated in a phospholipid membrane. This second subclass of phospholipid-based nanosensors, known as lipobead nanosensors, is more stable and less susceptible to biological degradation than liposome-based sensors. An additional advantage of these lipobead nanosensors over liposome nanosensors or PEBBLEs is that the fluorescent indicator dyes can be partially embedded in the phospholipid membrane, allowing the measurement of a much larger number of chemical species, because the analyte does not need to diffuse through a protective coating before interacting with the dye.[73,75]

CONCLUSION

With interest in nanotechnology and its practical use rising, the development of optical nanosensors for microscopic analyses has increased dramatically over the last decade. Since the development of the first fiber optic nanoprobes for near-field scanning optical microscopy, optical nanosensors have evolved into many different forms, each having its own distinct advantages for a particular type of analysis. Furthermore, these optical nanosensors and nanobiosensors have begun to demonstrate their ability to obtain reliable and useful measurements of chemicals species within cellular and even subcellular environments. Based on the rapid impact that optical nanosensor and nanobiosensor technologies have already had on cellular biology and biomedical diagnostics, future developments (i.e., smaller, less invasive sensors; more biocompatible sensors; etc.) should revolutionize the fields of healthcare and pharmaceutical development by providing a much greater understanding of basic cellular reaction pathways for various biological functions and diseases.

O

ACKNOWLEDGMENTS

This work was supported, in part, by the University of Maryland Baltimore County. In addition, I would like to acknowledge the generous assistance of Dr. Dimitra Stratis-Cullum with the graphics.

REFERENCES

1. Apuzzo, M.L.J.; Liu, C.Y. Things to come. Neurosurgery **2001**, *49* (4), 765–778.
2. Space mission for nanosensors. Futurist **2002**, *36* (6), 13.
3. Aylott, J.W. Optical nanosensors—An enabling technology for intracellular measurements. Analyst **2003**, *128* (4), 309–312.
4. Cui, Y.; Wei, Q.Q.; Park, H.K.; Lieber, C.M. Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. Science **2001**, *293* (5533), 1289–1292.
5. Cullum, B.M.; Vo-Dinh, T. The development of optical nanosensors for biological measurements. Trends Biotechnol. **2000**, *18* (9), 388–393.
6. Clark, H.A.; Hoyer, M.; Parus, S.; Philbert, M.A.; Kopelman, M. Optochemical nanosensors and subcellular applications in living cells. Mikrochim. Acta **1999**, *131* (1–2), 121–128.
7. Nice, E.C.; Catimel, B. Instrumental biosensors: New perspectives for the analysis of biomolecular interactions. Bioessays **1999**, *21* (4), 339–352.
8. Weetall, H.H. Chemical sensors and biosensors, update, what, where, when and how. Biosens. Bioelectron. **1999**, *14* (2), 237–242.
9. Tess, M.E.; Cox, J.A. Chemical and biochemical sensors based on advances in materials chemistry. J. Pharm. Biomed. Anal. **1999**, *19* (1–2), 55–68.
10. Braguglia, C.M. Biosensors: An outline of general principles and application. Chem. Biochem. Eng. Q. **1998**, *12* (4), 183–190.
11. Haugland, R.P. Indicators for Ca^{2+}, Mg^{2+}, Zn^{2+} and Other Metals. In *Handbook of Fluorescent Probes and Research Chemicals*, 6th Ed.; Spence, M.T.Z., Ed.; Molecular Probes, Inc.: Eugene, OR, 1996; 503–540.
12. Lu, J.Z.; Rosenzweig, Z. Nanoscale fluorescent sensors for intracellular analysis. Fresenius J. Anal. Chem. **2000**, *366* (6–7), 569–575.

13. Vo-Dinh, T.; Cullum, B.M. Biosensors and biochips: Advances in biological and medical diagnostics. Fresenius J. Anal. Chem. **2000**, *366* (6–7), 540–551.
14. Vo-Dinh, T.; Cullum, B.M.; Stokes, D.L. Nanosensors and biochips: Frontiers in biomolecular diagnostics. Sens. Actuators, B, Chem. **2001**, *74* (1–3), 2–11.
15. Tan, W.H.; Kopelman, R.; Barker, S.L.R.; Miller, M.T. Ultrasmall optical sensors for cellular measurements. Anal. Chem. **1999**, *71* (17), 606A–612A.
16. Cullum, B.M.; Vo-Dinh, T. Nanosensors: Development and Application of Nanosensors for Cellular Diagnostics. In *Biomedical Photonics Handbook,* 1st Ed.; Vo-Dinh, T., Ed.; CRC Press: Boca Raton, 2003.
17. Betzig, E.; Harootunian, A.; Lewis, A.; Isaacson, M. Near-field scanning optical microscopy (NSOM)—Investigation of radiation transmitted through subwavelength apertures. Biophys. J. **1985**, *47* (2), A407.
18. Betzig, E.; Lewis, A.; Harootunian, A.; Isaacson, M.; Kratschmer, E. Near-field scanning optical microscopy (NSOM)—Development and biophysical applications. Biophys. J. **1986**, *49* (1), 269–279.
19. Durig, U.; Pohl, D.W.; Rohner, F. Near-field optical-scanning microscopy. J. Appl. Phys. **1986**, *59* (10), 3318–3327.
20. Pohl, D.W. Scanning Near-Field Optical Microscopy. In *Advances in Optical and Electron Microscopy*; Mulvey, T., Ed.; Academic Press: London, 1984.
21. Teague, E.C. *Scanning Microscopy Technology and Applications*; Society of Photo-Optical Instrumentation Engineering: Bellingham, 1988; Vol. 897.
22. Betzig, E.; Trautman, J.K.; Harris, T.D.; Weiner, J.S.; Kostelak, R.L. Breaking the diffraction barrier—Optical microscopy on a nanometric scale. Science **1991**, *251* (5000), 1468–1470.
23. Betzig, E.; Chichester, R.J. Single molecules observed by near-field scanning optical microscopy. Science **1993**, *262* (5138), 1422–1425.
24. Betzig, E.; Isaacson, M.; Lewis, A. Collection mode near-field scanning optical microscopy. Appl. Phys. Lett. **1987**, *51* (25), 2088–2090.
25. Lieberman, K.; Harush, S.; Lewis, A.; Kopelman, R. A light-source smaller than the optical wavelength. Science **1990**, *247* (4938), 59–61.
26. Tan, W.H.; Shi, Z.Y.; Smith, S.; Birnbaum, D.; Kopelman, R. Submicrometer intracellular chemical optical fiber sensors. Science **1992**, *258* (5083), 778–781.
27. Tan, W.H.; Shi, Z.Y.; Kopelman, R. Development of submicron chemical fiber optic sensors. Anal. Chem. **1992**, *64* (23), 2985–2990.
28. Alarie, J.P.; VoDinh, T. Antibody-based submicron biosensor for benzo a pyrene DNA adduct. Polycycl. Aromat. Compd. **1996**, *8* (1), 45–52.
29. Cullum, B.M.; Griffin, G.D.; Miller, G.H.; Vo-Dinh, T. Intracellular measurements in mammary carcinoma cells using fiber-optic nanosensors. Anal. Biochem. **2000**, *277* (1), 25–32.
30. Hoffmann, P.; Dutoit, B.; Salathe. R.P. Comparison of mechanically drawn and protection layer chemically etched optical fiber tips. Ultramicroscopy **1995**, *61* (1–4), 165–170.
31. Turner, D.R. US patent #4,469,554. 1984.
32. Stockle, R.; Fokas, C.; Deckert. V.; Zenobi, R.; Sick, B.; Hecht, B.; Wild, U.P. High-quality near-field optical probes by tube etching. Appl. Phys. Lett. **1999**, *75* (2), 160–162.
33. Zeisel, D.; Deckert, V.; Zenobi, R.; Vo-Dinh, T. Near-field surface-enhanced Raman spectroscopy of dye molecules adsorbed on silver island films. Chem. Phys. Lett. **1998**, *283* (5–6), 381–385.
34. Deckert, V.; Zeisel, D.; Zenobi, R.; Vo-Dinh, T. Near-field surface enhanced Raman imaging of dye-labeled DNA with 100-nm resolution. Anal. Chem. **1998**, *70* (13), 2646–2650.
35. Munkholm, C.; Walt, D.R.; Milanovich, F.P. Preparation of CO_2 fiber optic chemical sensor. Abstr. Pap.—Am. Chem. Soc. **1987**, *193*, 183-ANYL.
36. Munkholm, C.; Parkinson, D.R.; Walt, D.R. Intramolecular fluorescence self-quenching of fluoresceinamine. J. Am. Chem. Soc. **1990**, *112* (7), 2608–2612.
37. Barker, S.L.R.; Thorsrud, B.A.; Kopelman, R. Nitrite- and chloride-selective fluorescent nanooptodes and in in vitro application to rat conceptuses. Anal. Chem. **1998**, *70* (1), 100–104.
38. Samuel, J.; Strinkovski, A.; Lieberman, K.; Ottolenghi, M.; Avnir, D.; Lewis, A. Miniaturization of organically doped sol–gel materials—A microns-size fluorescent pH sensor. Mater. Lett. **1994**, *21* (5–6), 431–434.
39. McCulloch, S.A.; Uttamchandani, D. IEE Proc., Optoelectron. **1995**, *144* (162).
40. Tan, W.H.; Shi, Z.Y.; Kopelman, R. Miniaturized fiber optic chemical sensors with fluorescent dye-doped polymers. Sens. Actuators. B, Chem. **1995**, *28* (2), 157–163.
41. Song, A.; Parus, S.; Kopelman, R. High-performance fiber optic pH microsensors for practical physiological measurements using a dual-emission sensitive dye. Anal. Chem. **1997**, *69* (5), 863–867.

42. Koronczi, I.; Reichert, J.; Heinzmann, G.; Ache, H.J. Development of a submicron optochemical potassium sensor with enhanced stability due to internal reference. Sens. Actuators, B, Chem. **1998**, *51* (1–3), 188–195.
43. Bui, J.D.; Zelles, T.; Lou, H.J.; Gallion, V.L.; Phillips, M.I.; Tan, W.H. Probing intracellular dynamics in living cells with near-field optics. J. Neurosci. Methods **1999**, *89* (1), 9–15.
44. Barker, S.L.R.; Kopelman, R. Development and cellular applications of fiber optic nitric oxide sensors based on a gold-adsorbed fluorophore. Anal. Chem. **1998**, *70* (23), 4902–4906.
45. Barker, S.L.R.; Kopelman, R.; Meyer, T.E.; Cusanovich, M.A. Fiber-optic nitric oxide-selective biosensors and nanosensors. Anal. Chem. **1998**, *70* (5), 971–976.
46. Cullum, B.M.; Vo-Dinh, T. Optical nanosensors and biological measurements. Biofutur **2000**, *2000* (205), A1–A6.
47. Vo-Dinh, T.; Griffin, G.D.; Alarie, J.P.; Cullum, B.M.; Sumpter, B.; Noid, D. Development of nanosensors and bioprobes. J. Nanopart. Res. **2000**, *2*, 17.
48. Vo-Dinh, T.; Alarie, J.P.; Cullum, B.M.; Griffin, G.D. Antibody-based nanoprobe for measurement of a fluorescent analyte in a single cell. Nat. Biotechnol. **2000**, *18* (7), 764–767.
49. Tyagi, S. Molecular beacons: Probes that fluoresce upon hybridization. Nat. Biotechnol. **1995**, *14* (3), 303–309.
50. Tyagi, S.; Kramer, F.R. Molecular beacons: Probes that fluoresce upon hybridization. Nat. Biotechnol. **1996**, *14* (3), 303–308.
51. Fang, X.H.; Li, J.W.J.; Perlette, J.; Tan, W.H.; Wang, K.M. Molecular beacons—Novel fluorescent probes. Anal. Chem. **2000**, *72* (23), 747A–753A.
52. Liu, X.J.; Tan, W.H. A fiber-optic evanescent wave DNA biosensor based on novel molecular beacons. Anal. Chem. **1999**, *71* (22), 5054–5059.
53. Liu, X.J.; Farmerie, W.; Schuster, S.; Tan, W.H. Molecular beacons for DNA biosensors with micrometer to submicrometer dimensions. Anal. Biochem. **2000**, *283* (1), 56–63.
54. Cordek, J.; Wang, X.W.; Tan, W.H. Direct immobilization of glutamate dehydrogenase on optical fiber probes for ultrasensitive glutamate detection. Anal. Chem. **1999**, *71* (8), 1529–1533.
55. Rosenzweig, Z.; Kopelman, R. Analytical properties and sensor size effects of a micrometer-sized optical fiber glucose biosensor. Anal. Chem. **1996**, *68* (8), 1408–1413.
56. Cullum, B.M.; Griffin, G.D.; Vo-Dinh, T. Nanosensors for analysis of a single cell. SPIE Proc. **2001**, *4254*, 35–40.
57. Li, H.; Chandras, N.; Cullum, B.M. Development and optimization of SERS-based immuno-nanosensor for single cell analyses. SPIE Proc. **2003**, *in press.*
58. Walt, D. Near-field array, (accessed July 2003). http://ase. tufts.edu/chemistry/walt/index.htm.
59. Chan, W.C.W.; Nie, S.M. Quantum dot bioconjugates for ultrasensitive nonisotopic detection. Science **1998**, *281* (5385), 2016–2018.
60. Lyon, W.A.; Nie, S.M. Single molecule methodologies for DNA analysis. Abstr. Pap.—Am. Chem. Soc. **1997**, *213*, 49-PHYS.
61. Taylor, J.R.; Fang, M.M.; Nie, S.M. Probing specific sequences on single DNA molecules with bioconjugated fluorescent nanoparticles. Anal. Chem. **2000**, *72* (9), 1979–1986.
62. Zhang, C.Y.; Ma, H.; Nie, S.M.; Ding, Y.; Jin, L.; Chen, D.Y. Quantum dot-labeled trichosanthin. Analyst **2000**, *125* (6), 1029–1031.
63. Zhang, C.Y.; Ma, H.; Ding, Y.; Jin, L.; Chen, D.Y.; Miao, Q.; Nie, S.M. Studies on quantum dots-labeled trichosanthin. Chem. J. Chin. Univ.—Chin. **2001**, *22* (1), 34–37.
64. Sasaki, K.; Shi, Z.Y.; Kopelman, R.; Masuhara, H. Three-dimensional pH microprobing with an optically-manipulated fluorescent particle. Chem. Lett. **1996**, *25* (2), 141–142.
65. Clark, H.A.; Hoyer, M.; Philbert, M.A.; Kopelman, R. Optical nanosensors for chemical analysis inside single living cells: 1. Fabrication, characterization, and methods for intracellular delivery of PEBBLE sensors. Anal. Chem. **1999**, *71* (21), 4831–4836.
66. Clark, H.A.; Kopelman, R.; Tjalkens, R.; Philbert, M.A. Optical nanosensors for chemical analysis inside single living cells: 2. Sensors for pH and calcium and the intracellular application of PEBBLE sensors. Anal. Chem. **1999**, *71* (21), 4837–4843.
67. Sumner, J.P.; Aylott, J.W.; Monson, E.; Kopelman, R. A fluorescent PEBBLE nanosensor for intracellular free zinc. Analyst **2002**, *127* (1), 11–16.
68. Clark, H.A.; Barker, S.L.R.; Brasuel, M.; Miller, M.T.; Monson, E.; Parus, S.; Shi, Z.Y.; Song, A.; Thorsrud, B.; Kopelman, R.; Ade, A.; Meixner, W.; Athey, B.; Hoyer, M.; Hill, D.; Lightle, R.; Philbert, M.A. Subcellular optochemical nanobiosensors: Probes encapsulated by biologically localised embedding (PEBBLEs). Sens. Actuators, B, Chem. **1998**, *51* (1–3), 12–16.
69. Xu, H.; Aylott, J.W.; Kopelman, R.; Miller, T.J.; Philbert, M.A. A real-time ratiometric method for

the determination of molecular oxygen inside living cells using sol–gel-based spherical optical nanosensors with applications to rat C6 glioma. Anal. Chem. **2001**, *73* (17), 4124–4133.

70. Xu, H.; Aylott, J.; Kopelman, R. Sol–gel pebble sensors for biochemical analysis inside living cells. Abstr. Pap.—Am. Chem. Soc. **2000**, *219*, 97-ANYL.
71. Anker, J.N.; Behrend, C.; Kopelman, R. Aspherical magnetically modulated optical nanoprobes (MagMOONs). J. Appl. Phys. **2003**, *93* (10), 6698–6700.
72. Anker, J.N.; Kopelman, R. Magnetically modulated optical nanoprobes. Appl. Phys. Lett. **2003**, *82* (7), 1102–1104.
73. Ji, J.; Rosenzweig, N.; Jones, I.; Rosenzweig, Z. Molecular oxygen-sensitive fluorescent lipobeads for intracellular oxygen measurements in murine macrophages. Anal. Chem. **2001**, *73* (15), 3521–3527.
74. Ji, J.; Rosenzweig, N.; Griffin, C.; Rosenzweig, Z. Synthesis and application of submicrometer fluorescence sensing particles for lysosomal pH measurements in murine macrophages. Anal. Chem. **2000**, *72* (15), 3497–3503.
75. McNamara, K.P.; Nguyen, T.; Dumitrascu, G.; Ji, J.; Rosenzweig, N.; Rosenzweig, Z. Synthesis, characterization, and application of fluorescence sensing lipobeads for intracellular pH measurements. Anal. Chem. **2001**, *73* (14), 3240–3246.
76. Nguyen, T.; McNamara, K.P.; Rosenzweig, Z. Optochemical sensing by immobilizing fluorophore-encapsulating liposomes in sol–gel thin films. Anal. Chim. Acta **1999**, *400*, 45–54.
77. McNamara, K.P.; Rosenzweig, N.; Rosenzweig, Z. Liposome-based optochemical nanosensors. Mikrochim. Acta **1999**, *131* (1–2), 57–64.
78. McNamara, K.P.; Rosenzweig, Z. Dye-encapsulating liposomes as fluorescence-based oxygen nanosensors. Anal. Chem. **1998**, *70* (22), 4853–4859.
79. McNamara, K.P.; Rosenzweig, Z. Nanosized phospholipid-coated particles as biological sensors. Abstr. Pap.—Am. Chem. Soc. **1999**, *218*, 26-ANYL.
80. DeCoster, M.A.; Nguyen, T.; Bazan, N.G.; Rosenzweig, Z. Adherence of lectin-coated micro lipobeads to cortical neurons in cell culture. FASEB J. **2002**, *16* (4), A37.

Ordered Vesicles at the Silicon–Water Interface

Duncan J. McGillivray
Oxford University, Oxford, United Kingdom

INTRODUCTION

The purpose of this article is to examine the stable, ordered near-surface phase of dichain cationic surfactants, which has been recently reported,[1] and to determine whether this phase can be unambiguously assigned as a vesicular phase. Such a phase would be a remarkable example of an aligned near-surface phase, extending hundreds of nanometers into solution, which exists in equilibrium with the bulk aqueous phase, and may explain the superspreading characteristics of these surfactants,[2] providing a reservoir of surfactant near the surface.

These surfactants are also widely used as fabric softeners and are the subjects of keen interest for pharmaceutical applications. They have been shown to form vesicles in dilute solutions. This is particularly significant as it extends the range of standard vesicle-forming compounds beyond phospholipids, long the gold standard of the field. Furthermore, the vesicles formed by these cationic surfactants are generally smaller than those formed from phospholipids, ranging from as small as 30 nm. These cationic vesicles are the subjects of biomedical research, where they are highly regarded as effective gene transfer agents in gene therapy.[3] Part of the success of this technique depends on the interaction between the cationic vesicle–DNA complex and the target cell membrane, which enables the DNA to cross undegraded,[4] and thus it is of interest to know how the surfactant vesicles behave near a surface.

VESICLES

The vesicle morphology is a common motif in surfactant self-arrangement. Favored by surfactants whose packing parameter falls in the typically lamellar range ($0.5 < pp < 1$), it often coexists in solution with other lamellar-type structures such as lamellar fragments or sponge phases. The precise morphology chosen for a given surfactant depends critically on the precise nature of the surfactant and its counterion and the solution conditions of the system.

The most well-studied and common vesicle-forming surfactants are phospholipids.[5] These vesicles have been used as cell-membrane analogs[6,7] and have a great deal in common with ordinary biological cells. They do not form spontaneously, requiring significant energy input to form typically through sonication, and produce large, metastable vesicles which can grow as large as micrometers. Phospholipid vesicles are not generally monodisperse and can be both unilamellar and multilamellar. Monodispersity is often achieved by forming vesicles using a fixed-size extrusion technique.

Phospholipid vesicles are not, however, found to be stable when interacting with a solid surface. This enables them to be widely used to deposit phospholipid bilayers on surfaces for use as model membranes, where the vesicle collapses on contact to produce a lamellar bilayer at the surface.[8–10]

The range of vesicle-forming surfactant systems has recently been enlarged through the use of combinations of oppositely charged surfactants,[11–14] nonionic surfactants,[15,16] and dichain cationic surfactants.[17–20] The vesicle-forming dichain cationic surfactants, which are the dialkyldimethylammonium bromide (DAB) family, are the focus of the work referred to as the basis of this article.

SURFACE MEASUREMENTS

Neutron reflectometry provides the unique ability to probe the surface and near-surface region of a solid–liquid interface *nonintrusively*, an important condition for the study of such fragile systems. Neutrons of 1–10 Å wavelengths, which are typically used for surface measurements, penetrate a single crystal of silicon with greater than 70% transmittance over a path length of around 15 cm and are then reflected from the solid–liquid interface (see Fig. 1 for the standard geometry of a solid–liquid experiment). Neutron reflectometry also gives information about the near-surface structure in solution. The neutron penetration length in D_2O is of the order of microns for 10-Å neutrons, although, experimentally, it is found that it is difficult to detect species which are highly disperse or rough below the surface.

The reflection experiment measures the number of neutrons reflected as a function of the wave vector change on reflection (normally referred to as the momentum

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120014162

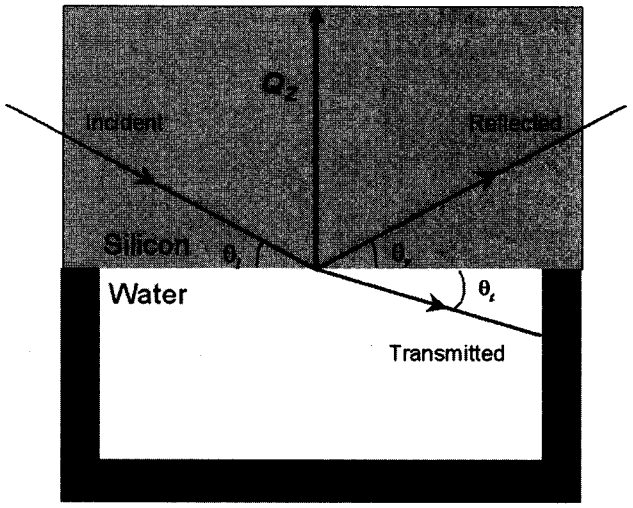


Fig. 1 The geometry of a specular neutron reflection experiment at the solid–liquid interface, where $\theta_i=\theta_r$ and Q_Z is the vertical momentum transfer. (*View this art in color at www. dekker.com.*)

transfer) Q_Z, which itself is a function of the wavelength of the neutrons and the angle of reflection

$$Q_Z = \frac{4\pi \sin\theta}{\lambda} \tag{1.1}$$

The reflection depends on the interaction of the neutrons with the nuclei present at the surface. The extent of nuclear interaction with a neutron is described by an experimentally well-determined quantity known as the coherent scattering length b_{coh}, which varies nonsystematically across the periodic table (see Table 1 for values of this parameter for a few biologically significant isotopes). In contrast to the X-ray scattering equivalent, it can be seen that hydrogen isotopes have large b_{coh}, which significantly takes a negative value for ^{1}H. This allows simple isotopic substitution to adjust the contrast of hydrogen-containing molecules; indeed, for water, the scattering length density (SLD) can range from -0.58 up to 6.35 Å^{-2}, enabling it to be contrast-matched to many common interfaces [SLD(Si)=2.07 Å^{-2}, SLD(Air)=0.0 Å^{-2}] so that the only scattering arises from material found at the interfacial region.

Table 1 Bound coherent scattering lengths (b_{coh}) for biologically significant isotopes

Element	b_{coh} (10^{-5} Å)	Element	b_{coh} (10^{-5}Å)
^{1}H	−3.74	^{12}C	6.65
^{2}H (D)	6.67	^{13}C	6.20
^{16}O	5.81	^{14}N	9.37

Source: Ref. [21].

Modeling

Data analysis is normally performed using optical matrix methods, in which the interface region is modeled as a series of homogenous layers. The layers are each characterized by a thickness [τ (Å)], SLD [ρ (Å^{-2})], and an interfacial roughness between adjacent layers [σ (Å^{-1})]. The exact reflectivity can be calculated from such a model; this can be compared with the data. and the model is then refined using least-squares minimization routines. For a more detailed discussion of the analysis of neutron reflectometry data in general, refer to Li et al.[22]

Off-Specular Scattering

If the interface between two regions is rough on a length scale appropriate to the wavelength of the scattering neutrons, then it is possible for some reflection to occur at angles that do not equal the incident angle. In this situation, the scattering, known as off-specular scattering, contains momentum transfer in both the perpendicular (Q_Z) and horizontal (Q_X) directions (see Fig. 2 below showing the geometry of the off-specular scattering experiment). It is therefore theoretically possible to determine, from the pattern of this scattering, information about the horizontal structure in a surface layer.

In practice, the determination of horizontal order is far from a routine procedure. Current instruments are not optimized for the detection of off-specular scattering, particularly at the small dispersive angles necessary for

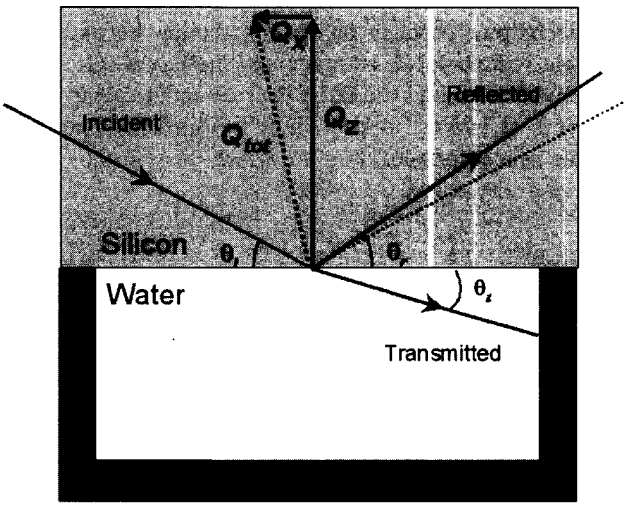


Fig. 2 The geometry of an off-specular reflection experiment. Here the angle of the reflected beam $\theta_r \neq \theta_i$, and there are two components to the total momentum transfer Q_{tot}, a vertical (Q_Z) and a horizontal (Q_X) component. (*View this art in color at www.dekker.com.*)

these DAB measurements, and there have been few measurements making use of this ability. The most notable exceptions are in the field of magnetic thin films, where the ability of polarized neutrons to interact with magnetic fields is employed.[23]

Bulk Solution Measurements

As an aid to modeling the surface conditions, bulk solution measurements are often used as a guide. Small-angle scattering is a well-established technique for determining bulk solution structures of surfactant solutions, which involves measuring the scattering of X-rays or neutrons from solutions in a transmission geometry. This technique is also nonperturbing, and neutron small-angle scattering shares many of the benefits of reflectometry in relation to the ability to alter contrast through the use of isotope labeling. X-ray small-angle scattering has the advantage of being able to make use of the extremely high flux available at modern synchrotron sources, enabling low-contrast (e.g., dilute) solutions to be measured, or measurements with high temporal resolution.

Cryo-transmission electron microscopy (cryo-TEM) is a much more recent technique, which involves taking a thin-film of a surfactant solution and flash-freezing it in near-freezing ethene. The freezing happens so rapidly that the water does not have time to crystallize and is considered to be much faster than what surfactant rearrangement would be, preserving the room-temperature surfactant structure.[24] However, it has been reported that the process of thinning the surfactant solution to an appropriate thickness for flash freezing (100–200 nm) may cause changes to the surfactant structure,[25] so that the evidence from cryo-TEM images must always be considered carefully.

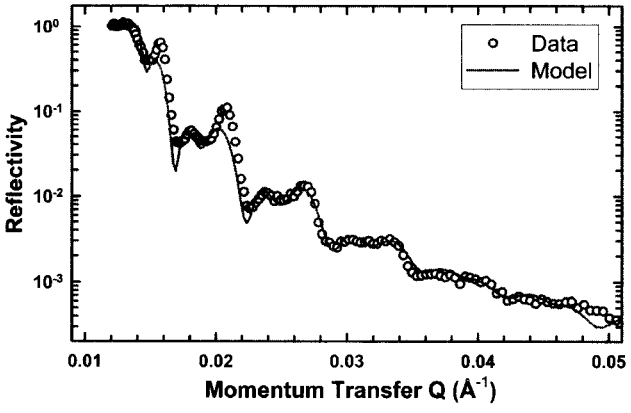


Fig. 3 Specular reflectivity profile and simple modeled fit to data for C12C12DAB 1% w/w solution in D_2O at 65°C. (*View this art in color at www.dekker.com.*)

Table 2 Parameters of the optical matrix fit to the scattering from a C12C12DAB 1% w/w solution in D_2O at 65°C

Layer	Silicon	SiO_2 layer	Void 1	Layer 1	Void 2	Layer 2
SLD/($Å^{-2}$)	2.07	3.4	−0.2	−0.2	−0.2	−0.2
Thickness/(Å)	n/a	12.0	690	170	630	165
Roughness/(Å)	5.0	5.0	40	30	40	30
Solvent/(%)	n/a	0	100	90	100	93

RESULTS

Fig. 3 shows a typical specular reflectivity profile for a sample of C12C12DAB (didodecyldimethylammonium bromide) at 1% w/w in solution, together with a simple modeled fit to the data. The parameters of this optical matrix model can be found in Table 2. Several features are immediately clear. The primary feature is the appearance of strong quasi-Bragg peaks in the specular reflectivity profile. These peaks are indicative of a periodic structure extending into solution perpendicular to the surface. The relatively poor definition of the peaks, and the appearance of subsidiary peaks, implies that the structure consists of a limited number of repeats into the solution. In this case, it appears that the pattern is adequately fit through the use of only two layers, but more detailed modeling[1] suggests up to four may be partially complete for this sample. Samples of other DABs show very similar behavior across the same temperature range.

The modeling implies a lamellar structure aligned with the surface, consisting of broad lamellar sheets separated by large solvent-filled voids of the order of 1000 Å. Such a pattern may be observed for several morphologies at the surface, which are illustrated in Fig. 4 below.

It is difficult to determine from a perpendicular density profile which of these options is the one occurring at the surface. Each could give rise to a periodic density modulation, as seen in the diagram (where the characteristic repeat distance is marked), and the vertical walls in the sponge and vesicle phases would contribute negligibly to the average SLD across the voids in the structure. A significant point that arises from the fitting is that the periodic repeat layer is of low density (approximately 90% solvent content) and much thicker than the typical bilayer thickness for a bilayer of a C12-surfactant sheet (ca. 20 Å). The interlayer spacing is also smaller than that calculated for a simple space-filling lamellar bilayer structure at similar concentrations (ca. 2000 Å), implying that this phase does not extend throughout the solution. This coincides with surface force measurements, which do

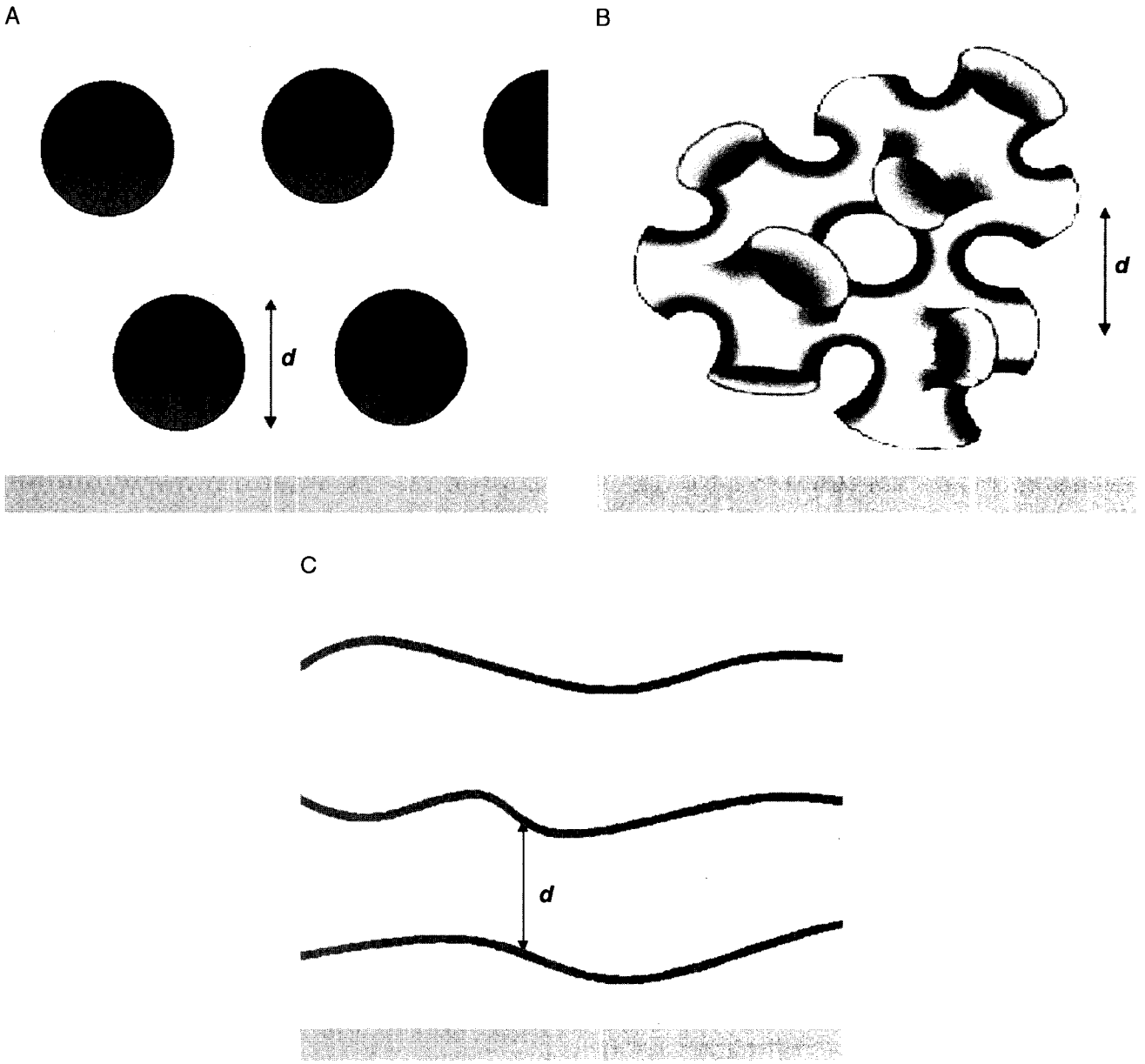


Fig. 4 Possible morphologies giving rise to lamellar layers at the surface. (A) Vesicles aligning at the surface; (B) a sponge phase; (C) lamellar sheets, with the characteristic repeat distance *d* marked. (*View this art in color at www.dekker.com.*)

not show any evidence of this phase,[26] implying that the layering detected in reflectometry is either too diffuse, too fragile, or both to be detected by a contact method.

These facts can be rationalized differently for each of possible morphologies. For a lamellar system, this observation would be consistent with an undulating bilayer, such that the density of the layer is averaged over the undulation period. This implies undulations of a 100-Å amplitude, which is reasonable for a lamellar sheet at this temperature unconstrained by near neighbors. Conversely, for a vesicular system, the low density and thick layers could be explained by a low horizontal packing density of vesicles, as would be expected given their surface charges and the fact that the curved surface of the vesicle is being projected into a planar averaging. Similar arguments could be used for the sponge phase.

Comparison with Bulk Structure

Some guidance can perhaps be gained from a knowledge of the bulk behavior of these surfactants. Phase diagrams

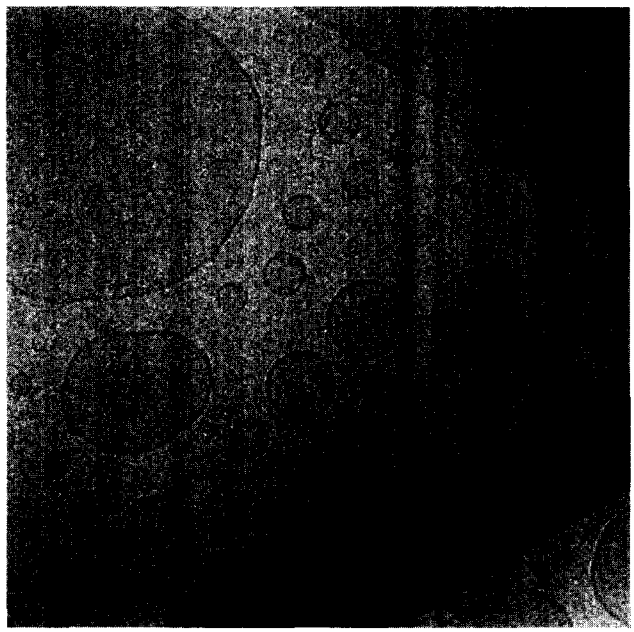
have been produced for dichain cationic surfactants of these types,[17,20,27] which show some general features in common. The first is that they all form a vesicular phase at low concentrations (approximately those of the systems described here, namely, of the order of 1% w/w). The stability of these vesicles has been questioned, it being suggested that they do not form spontaneously (requiring some form of energy input to form, such as sonication or shaking[14]) and will spontaneously collapse to a lamellar phase with time.[17] The solutions reported here were indeed sonicated on dissolution, but the structuring at the surface is stable over a period of hours, with no visible changes. Another marked difference between the bulk and the surface is that the vesicles imaged in the bulk phase (e.g., through freeze-fracture electron microscopy) are polydisperse, with radii of 1300±600 Å for a solution of C11C12DAB imaged by freeze-fracture TEM.[20] The degree of polydispersity would need to be much lower to give rise to the quasi-Bragg peaks seen in the specular reflectivity profiles measured for these systems.

However, it is also seen that either lamellar or sponge phases can coexist with this vesicle phase. Each of these phases is quite consistent with the simple geometric rationale of surfactant morphologies proposed by Tanford[28] and developed by Israelachvili et al.[29] as the packing parameter for the dichain surfactants, ca. 0.6 (varying ±0.01 depending on the chain length), lies in the middle of the range of lamellar-based structures. It appears that the bulk phase of the dichain surfactants is therefore quite complex and appears to be dependent on the precise history of the sample.

As a result, small-angle scattering and cryo-TEM experiments can be useful to determine the particular nature of these solution phases. Two representative cryo-TEM images are shown below (Fig. 5). These images clearly show that for the same sample, regions of vesicle and lamellar-like structures can be found. Indeed, other images also show that the vesicles and the lamellae were capable of coexisting in the same region. There is no evidence, however, of sponge phase structures found in any of the exposures. The lamellar structures imaged also show that the lamellae seen are not isolated lamellar sheets, but rather appear to be elongated microtubules—connected with a definite interior and exterior (as marked in Fig. 5b).

Although the lack of sponge phase observed in the cryo-TEM does not of itself exclude the possibility of sponge phase being found at the surface (particularly given the nature of cryo-TEM discussed above), it seems that this phase is the least likely of the postulated options. The inherently three-dimensional continuous structure of the sponge phase does not pack well into the highly asymmetric and constrained near-surface region and, as

A



B

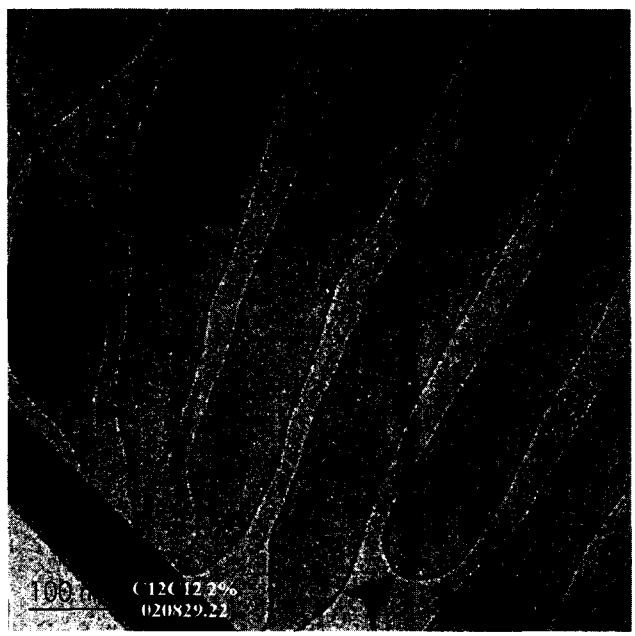


Fig. 5 Cryo-TEM images of C12C12DAB 2% w/w solution flash-frozen from ca. 25°C; A) showing polydisperse unilamellar vesicles in the solution; B) showing the "extended vesicle" or microtubules. The closure is marked with an arrow. Both images are taken from different areas of the same sample, scale as marked.

such, would need to show some form of transition from a free bulk sponge phase to a surface phase.

In any case, the small-angle scattering experiments provide evidence that the bulk phase behaves significantly differently to the near-surface phase—direct evidence for the limited nature of the surface-induced structural alterations. Although there is some evidence of a lamellar layering in the small-angle X-ray scattering, the structure is not as pronounced as that found at the surface (see Fig. 6 for a comparison of the structuring found for C12C12DAB, 1% w/w solution, for example) nor does bulk spacing (when observed) correspond to the surface spacing.

The differences between the bulk and the near-surface are also seen in the temperature response of the two systems. The interlayer spacing of the near-surface phase

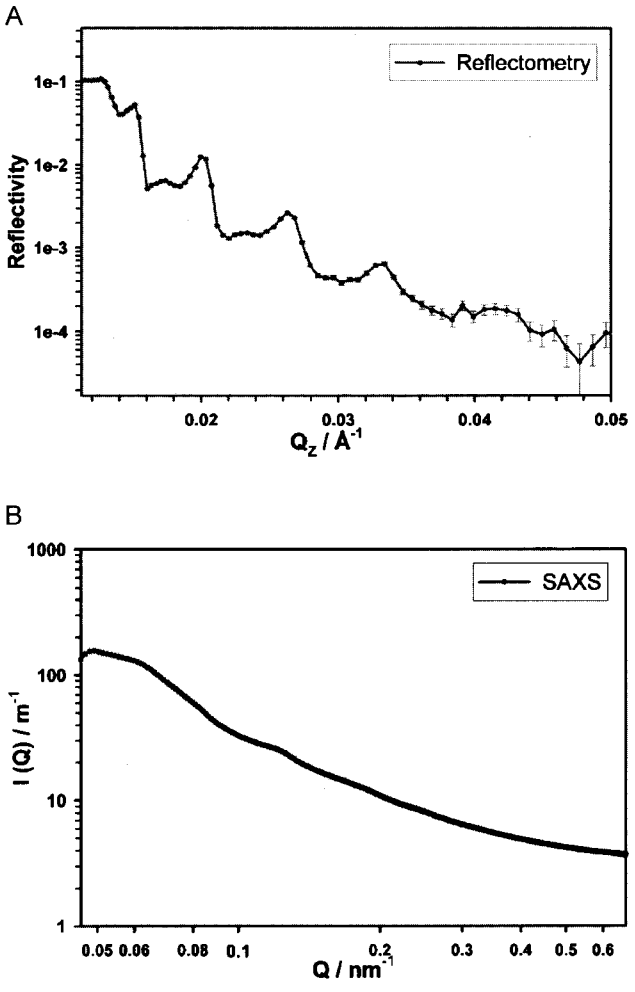


Fig. 6 Comparison of structuring for C12C12DAB 1% w/w solution at 45°C at the surface and in the bulk solution from A) reflectometry and B) small-angle scattering. (*View this art in color at www.dekker.com.*)

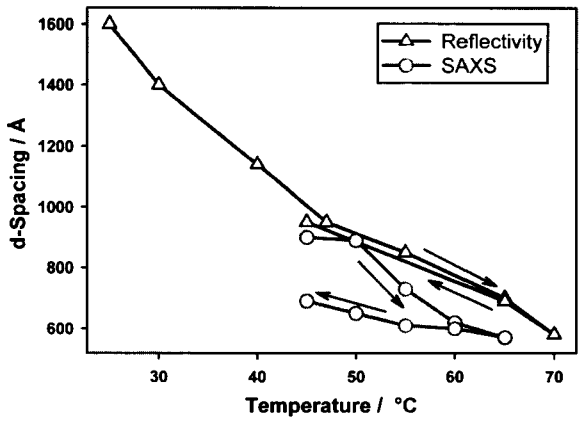


Fig. 7 The change in the d-spacing of the surfactant solution as determined from neutron reflectivity and small-angle scattering. The arrows indicate the direction of the temperature cycles in testing the reversibility of changes. (*View this art in color at www.dekker.com.*)

is highly sensitive to the temperature, generally decreasing with increasing temperature, as seen by the shift in the quasi-Bragg peaks in the scattering (Fig. 7). This type of temperature dependence at the surface has previously been reported for the lamellar phase of a 2% solution of NaAOT (sodium bis(2-ethylhexyl)sulfosuccinate) at the silicon–water interface, where the spacing changes from 270 Å at 5°C to 165 Å at 35°C.[30] In that system, the change in spacing was attributed to a change in an equilibrium between a micellar and a lamellar phase, where the surface structure was unambiguously lamellar extending thousands of layers into solution.

This change of spacing is reversible over at least a 30°C temperature range, which is also in contrast to the bulk solution, which is shown by small-angle scattering to be only broadly reversible (Fig. 7). This implies that the surface layer is in equilibrium at the surface and also that the time taken for the surface phase to reach this stable position is less than the equilibration time allowed after temperature changes (<1 hr).

The temperature dependence can again be justified in terms of both the vesicular and lamellar morphologies at the surface. As the temperature increases, the solvation shell of headgroups will decrease, and the degree of ionization of the bilayer will increase—effects which will tend to increase the rigidity of the bilayer and hence decrease the degree of fluctuation–repulsion. However, this effect will be offset by the increased thermal energy of the bilayer and the increased electrostatic repulsion of the more highly charged layers, and the net effect is difficult to predict.

In the case of the vesicular morphology, in particular, attention needs to be given to the manner of how exactly

the interlayer spacing might change. The direct implication is that the vesicles themselves change size, achieved either through the smooth aggregation of smaller vesicles to form larger vesicles (equivalently budding material on warming) or the loss of material to free surfactant as the solubility increases with temperature. Simple calculations for C12C12DAB show, for a unilamellar spherical surfactant vesicle, that the decrease in the volume of the bilayer in going from 1100 to 600 Å (between 40°C and 70°C) is approximately 70% of the original volume, equivalent to an increase of solubility of the order of 100 times. This is much greater than the predicted change in solubility, which means the former mechanism must be dominant to explain the behavior seen here for spherical vesicles.

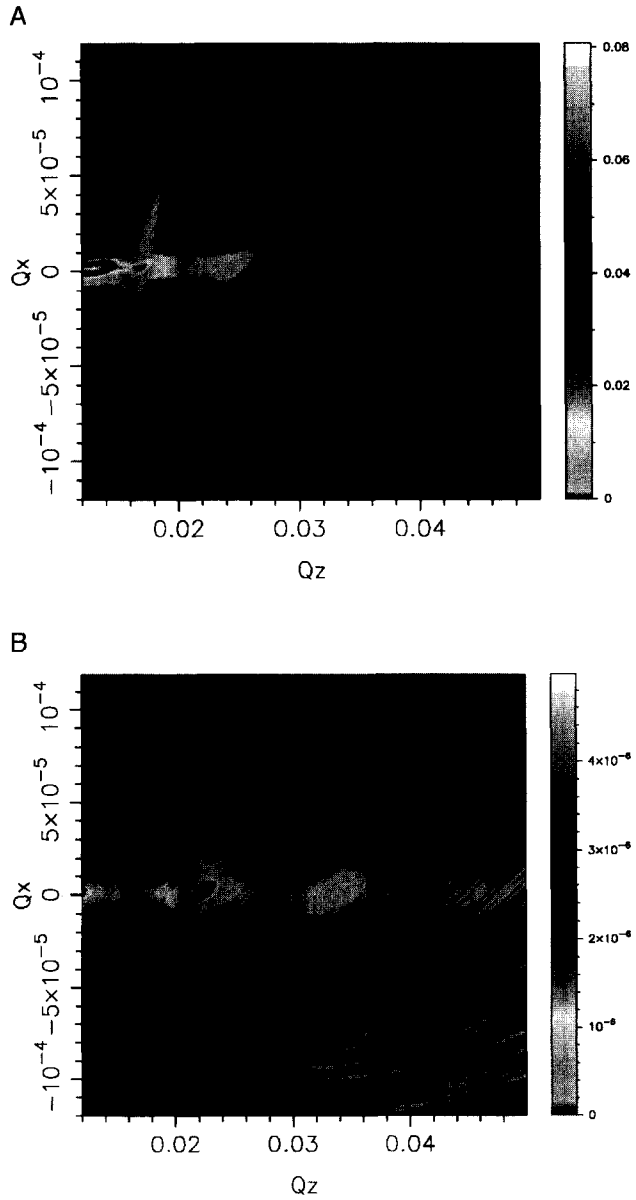


Fig. 8 Off-specular scattering from C12C12DAB 1% w/w in A) D_2O and B) CMSi, a mixture of H_2O and D_2O with a SLD equal to that of silicon crystal ($\rho = 2.07 \times 10^{-6}$ Å^{-2}). (*View this art in color at www.dekker.com.*)

Off-Specular Scattering

The samples exhibit significant off-specular scattering associated with the quasi-Bragg peaks, indicative of a degree of lateral ordering. Unfortunately, the length scale of the lateral resolution of the measurements is much larger than the equivalent vertical dimension (micrometers rather than angstroms), which means that any interference patterns on the same length scale as those found in the specular scattering, as might be envisioned from a horizontally ordered vesicular layer, would not be seen.

The scattering which is seen consists of bands of scattering across a range of Q_X at constant Q_Z (Fig. 8). The question of whether this is true near-surface scattering, or instead derives from the bulk, can be answered by reference to two features of the scattering—namely, the centering of the off-specular intensity decay on the specular ridge and the appearance of curvature in the bands as a result of refraction in the D_2O sample, which is absent in the sample contrast-matched to silicon (CMSi). This is also similar to the pattern seen for NaAOT and is believed to arise from correlated undulations between layers such as might be seen from the natural layer undulations.[31] These undulations might arise from either vesicular or lamellar membranes, but in the case of lamellar sheets, the large distance between the layers begs the question of the source of the correlation between the layers, failing connection between the sheets.

CONCLUSION

While it is clear that the system in question exhibits unusual surface properties, and that there is indeed a near-surface phase which differs from the bulk, it is extremely difficult to distinguish the possibilities that make up the structure. It is known that sponge, lamellar, and vesicular phases can exist in the solution, depending on the individual route of the preparation of the sample, but it is clear from small-angle scattering measurements that the surface is not exhibiting the same characteristics as the bulk. The exact surface density distribution, the temperature dependence of the structure, and the off-specular scattering all provide constraints on the possibilities at the

surface, but none of these specifically excludes any of the candidates. A sponge phase seems the least likely because of the complexity of the phase in the dimensionally constrained near-surface region and necessity for the phase to align with the surface. Simplicity suggests that the best candidate for the surface might be a lamellar phase; this, however, represents extremely large spacing for a lamellar phase, and it is difficult then to explain the correlated nature of the undulations seen from the off-specular scattering.

The third choice, vesicles, has its own complications—in particular, that the vesicles would be more monodisperse than the bulk solution, and that the mechanism for the change in the size of the vesicles with temperature is unclear. Perhaps the most reasonable compromise is suggested by the cryo-TEM images, which show the existence of an intermediary phase between vesicular and lamellar, equivalent to extended vesicles (microtubules). However, it does not seem possible to distinguish these possibilities based on the evidence now available.

ACKNOWLEDGMENTS

I would like to acknowledge the assistance in preparing this manuscript received from Dr. R.K. Thomas (University of Oxford, Oxford, United Kingdom) and the many useful conversations with Prof. J. Penfold (ISIS Facility, Rutherford-Appleton Lab, Didcot, United Kingdom) and Prof. A. Rennie (Studsvik, Uppsala University, Sweden). I would also like to acknowledge the assistance of Hanna Vacklin (University of Oxford) and Dr. Gunnel Karlsson at the University of Lund, Sweden, for assistance in obtaining the cryo-TEM images, and the Rhodes Trust, Oxford, for financial support through a scholarship.

REFERENCES

1. McGillivray, D.J.; Thomas, R.K.; Rennie, A.R.; Penfold, J.; Sivia, D.S. Ordered structures of di-chain cationic surfactants at interfaces. Langmuir **2003**, *19* (19), 7719–7726.
2. Svitova, T.F.; Hill, R.M.; Radke, C.J. Spreading of aqueous dimethyldidodecylammonium bromide surfactant droplets over liquid hydrocarbon substrates. Langmuir **1999**, *15* (21), 7392–7402.
3. Miller, A.D. Cationic liposomes for gene therapy. Angew. Chem., Int. Ed. **1998**, *37* (13–14), 1769–1785.
4. Templeton, N.S. Liposomal delivery of nucleic acids in vivo. DNA Cell Biol. **2002**, *21* (12), 857–867.
5. *Liposomes : A Practical Approach*; New, R.R.C., Ed.; Oxford University Press: Oxford, 1990.
6. Fendler, J.H. *Membrane Mimetic Chemistry: Characterizations and Applications of Micelles, Microemulsions, Monolayers, Bilayers, Vesicles, Host–Guest Systems, and Polyions*; Wiley: New York, 1982.
7. *Medical Applications of Liposomes*; Lasic, D.D., Papahadjopoulos, D., Eds.; Elsevier: Oxford, 1998.
8. Charitat, T.; Bellet-Amalric, E.; Fragneto, G.; Graner, F. Adsorbed and free lipid bilayers at the solid–liquid interface. Eur. Phys. J., B **1999**, *8* (4), 583–593.
9. Rapuano, R.; Carmona-Ribeiro, A.M. Supported bilayers on silica. J. Colloid Interface Sci. **2000**, *226* (2), 299–307.
10. Carmona-Ribeiro, A.M. Bilayer vesicles and liposomes as interface agents. Chem. Soc. Rev. **2001**, *30* (4), 241–247.
11. Bergstrom, M.; Pedersen, J.S. Small-angle neutron scattering (SANS) study of aggregates formed from aqueous mixtures of sodium dodecyl sulfate (SDS) and dodecyltrimethylammonium bromide (DTAB). Langmuir **1998**, *14* (14), 3754–3761.
12. Bergstrom, M.; Pedersen, J.S.; Schurtenberger, P.; Egelhaaf, S.U. Small-angle neutron scattering (SANS) study of vesicles and lamellar sheets formed from mixtures of an anionic and a cationic surfactant. J. Phys. Chem., B **1999**, *103* (45), 9888–9897.
13. Horbaschek, K.; Hoffmann, H.; Hao, J.C. Classic L-alpha phases as opposed to vesicle phases in cationic–anionic surfactant mixtures. J. Phys. Chem., B **2000**, *104* (13), 2781–2784.
14. Marques, E.F. Size and stability of catanionic vesicles: Effects of formation path, sonication, and aging. Langmuir **2000**, *16* (11), 4798–4807.
15. Ma, G.; Barlow, D.J.; Lawrence, M.J.; Heenan, R.K.; Timmins, P. Small-angle neutron-scattering studies of nonionic surfactant vesicles. J. Phys. Chem., B **2000**, *104* (39), 9081–9085.
16. Le, T.D.; Olsson, U.; Mortensen, K. Packing states of multilamellar vesicles in a nonionic surfactant system. Phys. Chem. Chem. Phys. **2001**, *3* (7), 1310–1316.
17. Dubois, M.; Zemb, T. Phase-behavior and scattering of double-chain surfactants in diluted aqueous-solutions. Langmuir **1991**, *7* (7), 1352–1360.
18. Caboi, F.; Monduzzi, M. Didodecyldimethylammonium bromide vesicles and lamellar liquid crystal. A multinuclear NMR and optical microscopy study. Langmuir **1996**, *12* (15), 3548–3556.
19. Marques, E.F.; Regev, O.; Khan, A.; Miguel, M.D.; Lindman, B. Vesicle formation and general phase

behavior in the catanionic mixture SDS-DDAB-water. The cationic-rich side. J. Phys. Chem., B **1999**, *103* (39), 8353–8363.
20. Haas, S.; Hoffmann, H.; Thunig, C.; Hoinkis, E. Phase and aggregation behaviour of double-chain cationic surfactants from the class of *N*-alkyl-*N*-alkyl′-*N*, *N*-dimethylammonium bromide surfactants. Colloid Polym. Sci. **1999**, *277* (9), 856–867.
21. Koester, L.; Rauch, H.; Seymann, E. Neutron scattering lengths: A survey of experimental data and methods. Atom. Data Nucl. Data **1991**, *49* (1), 65–120.
22. Li, Z.X.; Thirtle, P.N.; Weller, A.; Thomas, R.K.; Penfold, J.; Webster, J.R.P.; Rennie, A.R. Some improvements and extensions of the application of specular neutron reflection to the study of interfaces. Physica, B **1998**, *248*, 171–183.
23. Toperverg, B.P. Specular reflection and off-specular scattering of polarized neutrons. Physica, B **2001**, *297* (1–4), 160–168.
24. Almgren, M.; Edwards, K.; Karlsson, G. Cryo transmission electron microscopy of liposomes and related structures. Colloids Surf., A **2000**, *174* (1–2), 3–21.
25. Danino, D.; Talmon, Y.; Zana, R. Cryo-TEM of thread-like micelles: On-the-grid microstructural transformations induced during specimen preparation. Colloids Surf., A **2000**, *169* (1–3), 67–73.
26. Briscoe, W. University of Oxford, unpublished results.
27. Svitova, T.F.; Smirnova, Y.P.; Pisarev, S.A.; Berezina, N.A. Self-assembly in double-tailed surfactants in dilute aqueous-solutions. Colloids Surf., A **1995**, *98* (1–2), 107–115.
28. Tanford, C. Micelle shape and size. J. Phys. Chem. **1972**, *76* (21), 3020–3024.
29. Israelachvili, J.N.; Mitchell, D.J.; Ninham, B.W. Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. J. Chem. Soc., Faraday Trans. II **1976**, *72*, 1525–1568.
30. Li, Z.X.; Weller, A.; Thomas, R.K.; Rennie, A.R.; Webster, J.R.P.; Penfold, J.; Heenan, R.K.; Cubitt, R. Adsorption of the lamellar phase of aerosol-OT at the solid/liquid and air/liquid interfaces. J. Phys. Chem., B **1999**, *103* (49), 10800–10806.
31. Li, Z.X.; Lu, J.R.; Thomas, R.K.; Weller, A.; Penfold, J.; Webster, J.R.P.; Sivia, D.S.; Rennie, A.R. Conformal roughness in the adsorbed lamellar phase of aerosol-OT at the air–water and liquid–solid interfaces. Langmuir **2001**, *17* (19), 5858–5864.

Organofullerenes in Water

Eiichi Nakamura
Hiroyuki Isobe
The University of Tokyo, Tokyo, Japan

INTRODUCTION

Fullerenes[1–4] are entirely insoluble in water and by themselves do not show appreciable biological activity. However, suitable chemical modification allows the molecules to dissolve in aqueous solution to exhibit significant activity. In 1993, three papers reported simultaneously on the biological activities of fullerene carboxylic acids **1** and **2** (Fig. 1).[5–8] Studies ensued afterward revealed various interesting biological activities of organofullerenes owing to their photoactivity, radical quenching activity, and hydrophobicity.[9–11] The interests in biology stimulated the studies on the behavior of fullerenes in water and led to new discoveries on the formation of one- to three-dimensional supramolecular objects, including vesicle, rod, globule, membrane, and linear assemblies. Several fundamental properties of fullerenes form the background of such interest in fullerenes in water: extremely high hydrophobicity, high cohesive force between fullerene molecules,[12,13] photoactivity,[14] ability to accept and release electrons,[15] and relatively high reactivity that allows structural modifications.[16,17] This article highlights the behaviors of organofunctionalized fullerenes in water.

BIOLOGICAL EFFECTS OF WATER-SOLUBLE FULLERENES

An obvious obstacle for the studies of fullerenes in water is the lack of solubility of fullerenes in water.[18–22] One approach to solve the issue utilizes a solubilizing agent such as polyvinylpyridone[23] or cyclodextrin,[24] and the other relies on introduction of hydrophilic function(s) by chemical modification of the molecule. Attachment of amino or hydroxyl groups[25] is straightforward but affords an inseparable mixture comprising many products of similar properties. Carbon–carbon bond forming modifications of fullerenes is more selective. The compounds **1** and **2** represent the compounds in this category (Fig. 1).

Nakamura et al.[26] synthesized the compound **1** and discovered that the compound shows significant biological activities.[5] The compound **1** (and later its C_{70} analog)[27] suppresses growth of mammalian cells, inhibits activity of various enzymes (including HIV protease, which they found later; Fig. 3),[28] and cuts DNA upon visible light irradiation. Later studies indicated that the cleavage is caused by active oxygen species generated by fullerene-sensitized activation of molecular oxygen in the medium (Fig. 2).[29,30] The carboxylic acid **1** does not bind to DNA.[31]

Papers from 1993[6–8] described inhibition of HIV protease by the compound **2**. The dicarboxylic acid **2** acts as a reversible inhibitor of the enzyme through binding to the cleft-like active site of the enzyme in competition with the natural peptide substrate. This shape-recognition binding is illustrated in Fig. 3 for the carboxylic acid **1**.[28] A dendritic variation of the fullerene inhibitor is studied for treatment of drug-resistant HIV.[32,33]

Phamacokinetics and Toxicity of Water-Soluble Fullerenes

Toxicity is the primary concern for people working on fullerenes on the bench. Although earlier studies on C_{60} itself suggested low toxicity,[34–36] it was unclear if water-soluble fullerenes are also innocuous molecules. For the studies of the pharmacokinetic behavior of water-soluble organofullerenes, Nakamura et al.[37–40] synthesized the compounds **1** and **1-C14** starting with an alkoxy-substituted methylenecyclopropane by the use of radiolabeled trimethylenemethane chemistry as shown in Fig. 4.

The radiolabeled water-soluble fullerene **1-C14** quickly migrates through the body and accumulates mainly in the liver after a few hours. Importantly, the study also suggested that the fullerene molecule can pass through the blood–brain barrier, which was later confirmed by others.[41,42] The behavior stands in sharp contrast to the parent fullerene, which was absorbed by serum protein in blood quickly after injection.[34–36] The pharmacokinetic studies have shown that organofullerenes are excreted either slowly or rapidly depending on the substituents, which may be suitably modified under the capacity of medicinal chemists.[43]

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013596

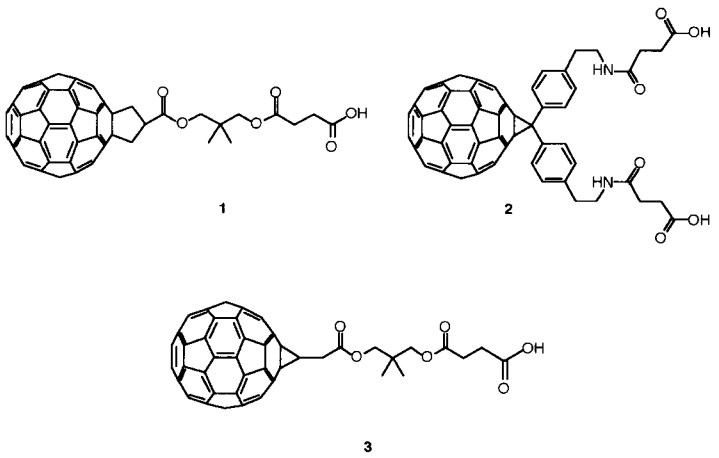


Fig. 1 Monocarboxylic and dicarboxylic acids **1** and **2**.

The toxicity studies on the carboxylic acid **3** showed that the molecule does not show serious acute toxicity. Thus upon injection of up to 500 mg/kg quantity of the compound **3** to female mice, all mice survived for 1 week. Later studies[44–46] also indicated that the fullerene molecules do not show appreciable short-term toxicity. Long-term toxicity has so far not been reported, and little has yet been known for degradation of fullerenes in the environment.[47,48]

Selective Binding to DNA

The initial discoveries made for the two simple molecules **1** and **2** quickly led to the studies on further application using more complex molecules. A fullerene molecule was connected to a 14-mer oligonucleotide that is complementary to single and double helix. The fullerene/DNA conjugate **4** was found to bind to the target DNA site (Fig. 5),[49] cleaving the guanine bases in the position expected on the basis of complementary of triple helix formation. A similar result was reported for a fullerene conjugate bearing a 31-mer DNA recognition site.[50] The binding ability of such DNA/fullerene conjugates has been assessed.[51] Fullerene molecules bearing an intercalating group have been shown also to bind to double-strand DNA.[52]

One application of the above study is the photodynamic therapy of cancer. A C_{60} derivative bearing polyether side chains was found to shrink skin cancer in mice.[53] In vivo ultrasound irradiation of fullerene-containing tissue may also suppress tumor growth.[54,55] Inactivation of virus has also been reported.[56] The actual biological targets of these approaches still remain unclear, although DNA may be a possible candidate.

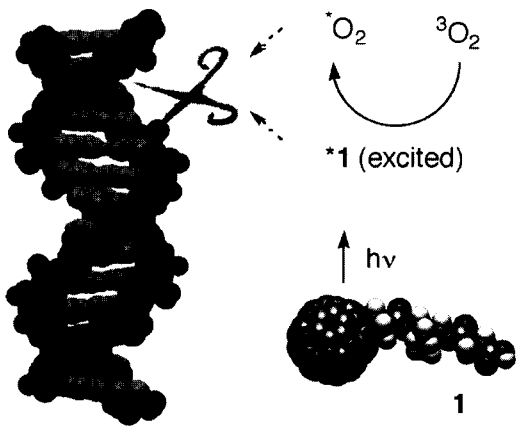


Fig. 2 DNA cleavage by fullerene carboxylic acid **1**. (*View this art in color at www.dekker.com.*)

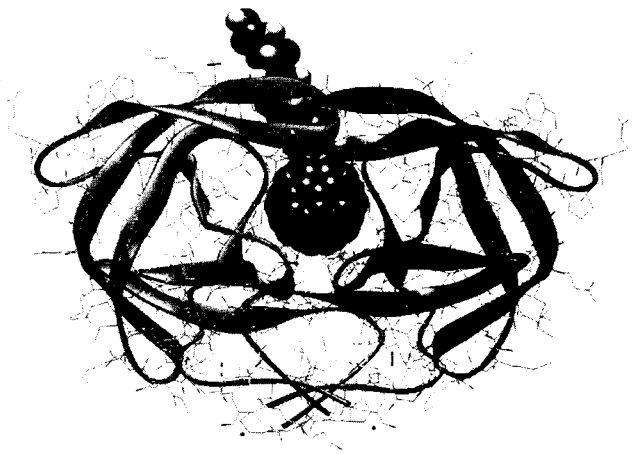


Fig. 3 A three-dimensional model of the complex of **1** with HIV protease. (*View this art in color at www.dekker.com.*)

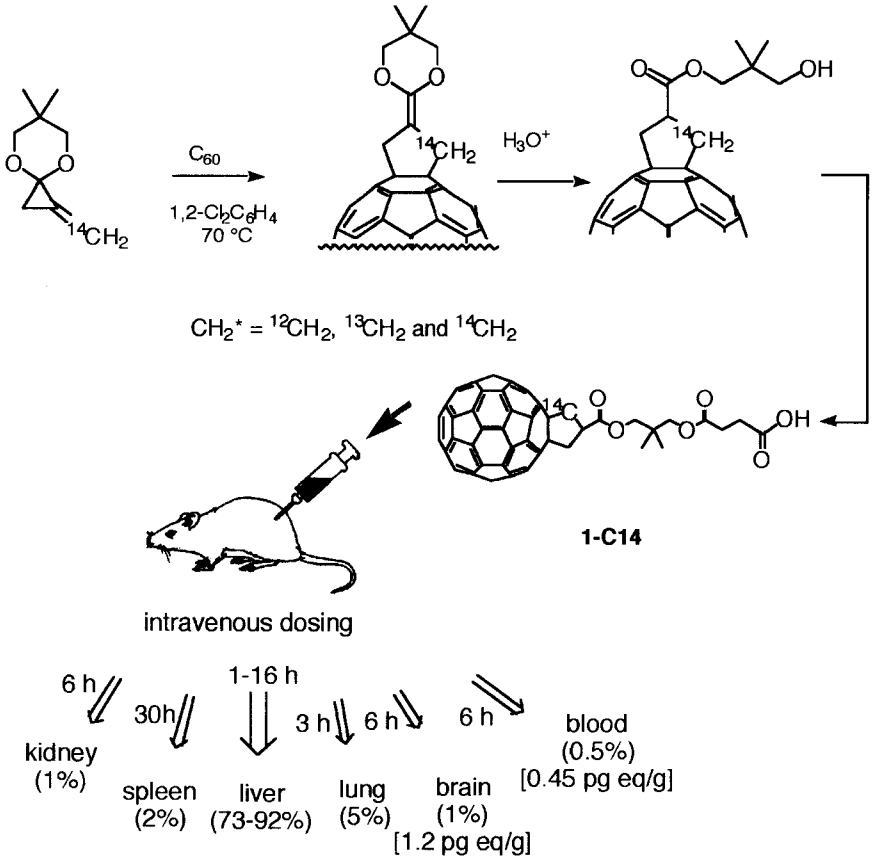


Fig. 4 Phamacokinetics of radiolabeled fullerene carboxylic acid **1-C14**.

Gene Delivery

Delivery of gene into living cells through encapsulation of the gene in a "vector" is a fundamental technology in modern biology.[57] This process called "transfection" (transfer+infection) allows the production of the protein encoded in the delivered DNA within the cell. Recent studies indicated that suitably modified fullerene molecules are effective for the gene delivery.

A series of fullerenes bearing amino side chain(s) has been synthesized and was examined for their DNA binding ability as well as their transfection capability.[58]

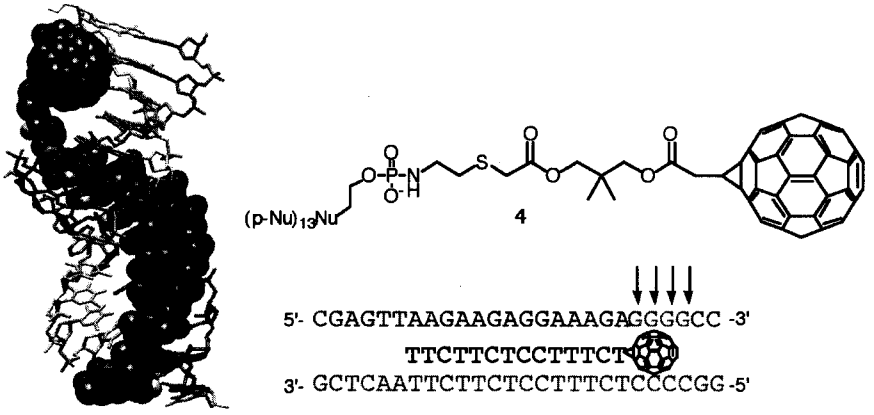


Fig. 5 Triple helix forming fullerene/DNA conjugate **4** (CPK model), its three-dimensional binding model, and site-selective photocleavage of DNA (as indicated by arrows). (*View this art in color at www.dekker.com.*)

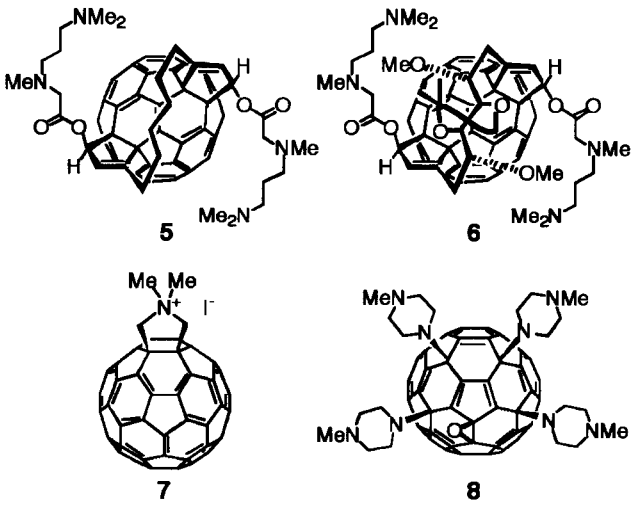


Chart 1 Representative DNA binding fullerenes examined for transfection.

The C_2-chiral two-handed fullerene **5** was synthesized by double [3+2] cycloaddition of a bis-cyclopropenone acetal that took place with complete regioselectivity to give the desired compound.[59]

Some representative examples of the DNA-binding fullerenes (**5–8**) are shown in Chart 1.[58,60–63] As a solution of a 4.5-kbp plasmid DNA with a reporter gene was mixed together with a solution of the two-handed tetraminofullerene **5** in buffer, particles of 0.5–3 μm diameter form as observed by optical microscope (Fig. 6). When this mixture was incubated with dividing mammalian cells, the dark-colored fullerene/DNA particles were taken into the cells, and protein production occurred. Screening of various cell lines indicated that the efficiency

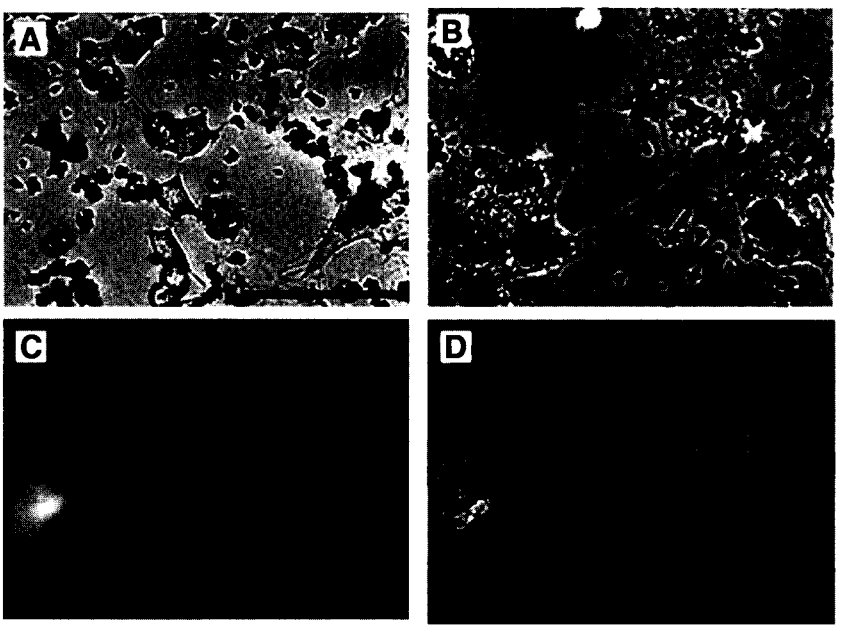


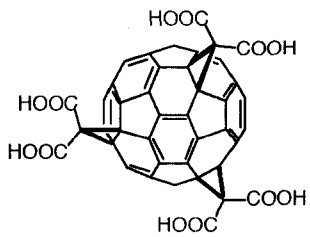
Fig. 6 Delivery of green fluorescent protein (GFP) plasmid into COS-1 cells with **5**. Bars show 100 μm. (A) Differential interference contrast microscopy (DIC) micrograph of COS-1 cells and small black particles (1–10 μm) of the fullerene/DNA complex after 1-hr transfection time. (B) DIC micrograph of COS-1 cells after 2-day incubation time. (C) Fluorescence image of the area same as (B). (D) Superimposed photograph of (B) and (C). Round black clumps of 0.5–3 μm size observed in the fluorescent cells contain fullerene materials. (Adapted from Ref. [31].) (*View this art in color at www.dekker.com.*)

of this first-generation fullerene vector is comparable to or better than the commercially available lipid-based vectors.

When a double-strand DNA is treated with an aminofullerene **7**, the fullerene molecules stick to the DNA strand without much changing the linear outline of the DNA as studied by TEM.[62,63] In spite of its binding ability, aminofullerene **7** is not effective for gene delivery.

Neuron Protection

In 1997, it was reported that the fullerene hexacarboxylic acid **9** acts as a drug for neurodegenerative diseases such as Alzheimer's, Parkinson's, or amyotrophic lateral sclerosis (ALS).[41] The activity is considered to arise from trapping of hydroxy radicals by the fullerene molecule.[64] The finding is currently developed for clinical trials.



9

Endohedral Metal Complexes

Metal-containing fullerenes are attracting people's attention for their clinical applications. Biodistribution studies of $^{166}Ho_x@C_{82}(OH)_y$ showed the uptake of the polyhydroxylated fullerene by bone without clearance and may be used as in vivo radiotracer.[42] The use of metallofullerenes as a magnetic resonance imaging (MRI) contrast agent is another area of extensive current studies. $Gd@C_{82}(OH)_n$ was found to be a much better contrast agent than the conventional agents (i.e., a nitrogen-ligand coordinated to the metal).[65,66] The metal-containing fullerene accelerates spin relaxation (T_1) of water protons that surrounds the fullerene molecule. The water proton relaxivity R_1 is much higher (20-folds) than that of the MRI contrast agent currently in use (e.g., Magnevist®, gadolinium-diethylenetri-aminepentaacetic acid). Polycarboxylic acid derivatives of $Gd@C_{82}$ also serve as an excellent contrast agent.[67]

MOLECULAR ASSEMBLY OF WATER-SOLUBLE FULLERENES

Many of "water-soluble" fullerenes do not simply dissolve in water, but form various intriguing supramolecular structures, some of which have been unknown for conventional surfactants and lipid molecules. Various new discoveries have been made in recent years with respect to the formation of one- to three-dimensional supramolecular objects, including vesicle, rod, globule, membrane, and linear assemblies.

Langmuir–Blodgett Film

Formation of molecular layers of fullerene and their derivatives has attracted the attention of scientists for a long time, yet earlier attempts to make uniform monolayer of C_{60} were rather unsatisfactory.[68] The amphiphilic fullerene carboxylic acid **1** was reported in 1995 to form a uniform Langmuir–Blodgett (LB) film at the water–air interface.[69] The average area per molecule determined by the π-A plot was 0.78 nm, and the distance between the neighboring C_{60} was 0.95 nm, which is close to (but slightly smaller than) the distance of 1.0 nm determined in the crystal structure of C_{60}. Atomic force microscopic (AFM) analysis indicated that the surface contains far less defects than the LB film made from unmodified C_{60} under similar conditions (Fig. 7).

Nanorod and Nanoparticles

Rod-like assembly of fullerene forms when a DMSO solution of a cationic amphiphile **7** is dispersed in benzene.[62,63] The nanorods measure a diameter of 14–120 nm and a length of over 70 μm. When the dispersion is sonicated, vesicles of various sizes and shapes form. The thickness of the vesicle wall (3–6 nm) suggests a multilamellar structure.[70] Georgakilas et al.[71] also reported that cationic amphiphiles form nanospheres and nanorods.

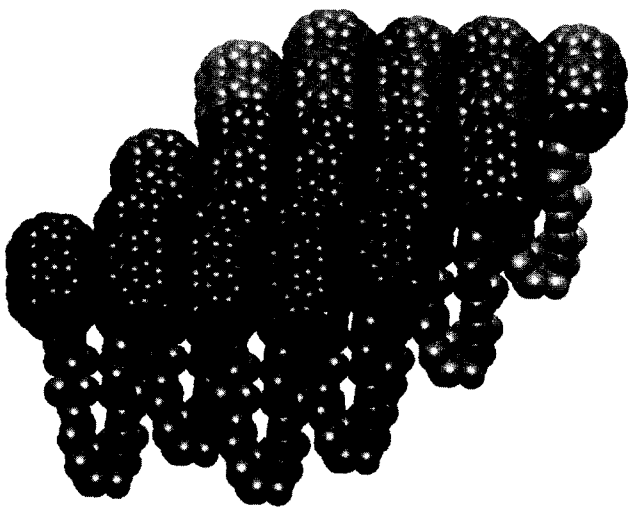
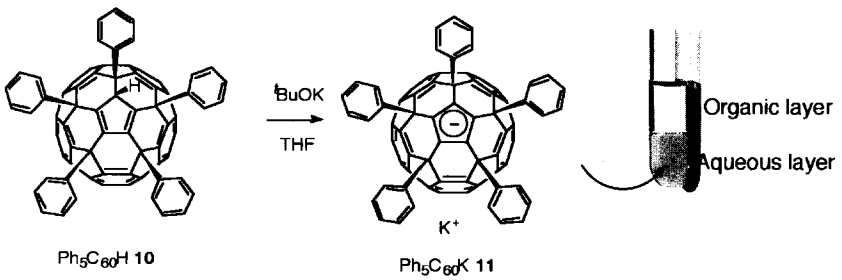


Fig. 7 Molecular model of LB film of **1**. (*View this art in color at www.dekker.com.*)



Scheme 1 (*View this art in color at www.dekker.com.*)

Bilayer Vesicles

$Ph_5C_{60}H$ (**10**) is an interesting hydrocarbon[72–74] in that its anion, $Ph_5C_{60}K$ (**11**), dissolves freely in water (Scheme I). It dissolves in water not as unimer, but as spherical aggregates.[75] Atomic force microscopy and laser light scattering (LLS) studies of the solution of **11** indicated that the solution contains bilayer vesicles and that the size distribution is quite narrow.[76] Thus upon slow addition of a THF solution to water, $Ph_5C_{60}K$ forms spherical bilayer aggregates of 34-nm average diameter composed of about 13,000 fullerene molecules as determined by LLS (Fig. 8). This ability of $Ph_5C_{60}K$ to form vesicles in water is the most intriguing. The Ph_5C_{60} anion amphiphile is noteworthy for its virtually all-sp^2 carbon framework, high structural rigidity, spherical hydrophobic moiety, and the lack of heteroatom hydrophilic sites.[77] In addition, the fullerene moiety is unique because it is highly hydrophobic but not lipophilic.

Penta-arylated C_{60} bears large aryl groups to form a cavity that can recognize another C_{60} molecule. Such a molecule **12** looks like a badminton shuttlecock and stacks with each other to form a one-dimensional array of both in crystals and in liquid crystals (Fig. 9).[78] One may expect that, with suitable modifications, such "nanoshuttlecock" would also be useful for the formation of supramolecular structures in aqueous solutions.

DNA/Fullerene Nanoparticles

The aforementioned experiments of gene delivery demonstrated the ability of the fullerene **5** to condense and to release DNA, and the molecular nature of the DNA condensation/release process was probed in test-tube experiments with AFM.[79] When half an equivalent/base pair of the fullerene **5** was mixed with 4-kbp plasmid

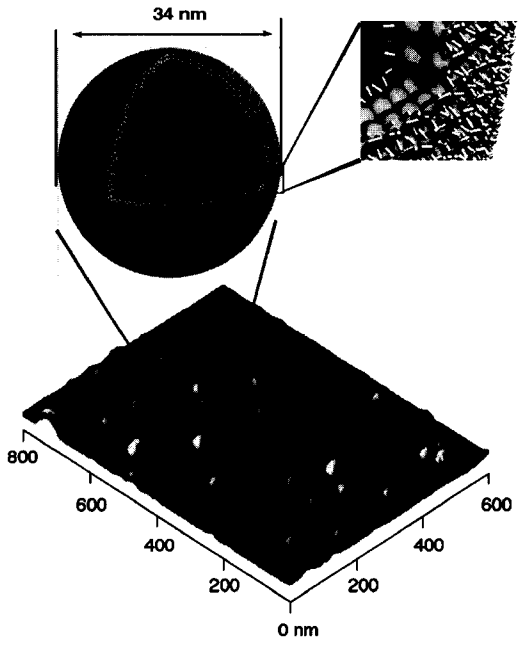


Fig. 8 Atomic force microscopic image of bilayer vesicle of **11** and the molecular model obtained by LLS study. (Adapted from Ref. [76].) (*View this art in color at www.dekker.com.*)

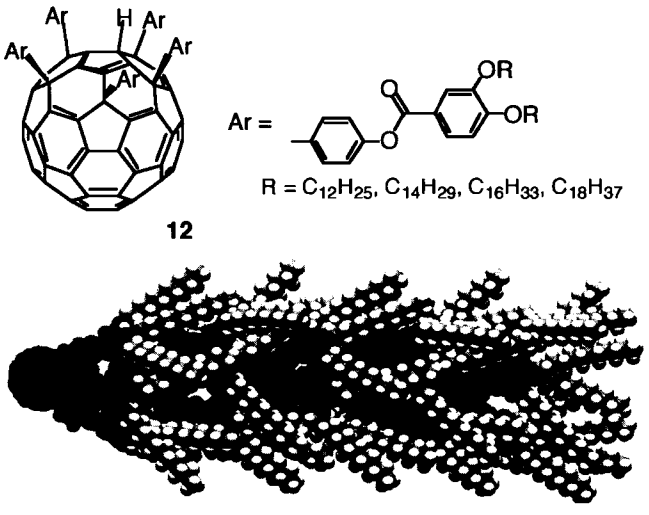


Fig. 9 Nanoshuttlecock molecule **12** and a three-dimensional model of the stack-of-five molecules. (*View this art in color at www.dekker.com.*)

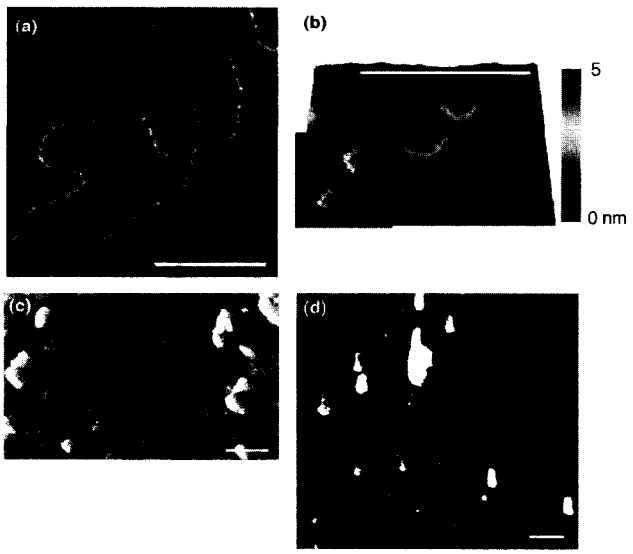


Fig. 10 Condensation of 4-kbp plasmid DNA (a) into globules with increasing amount of **5** (b and c) and its release upon removal of fullerene (d). (Adapted from Ref. [79].) (*View this art in color at www. dekker.com.*)

DNA (Fig. 10a), the DNA molecule was condensed into a globular object seen in Fig. 10b as a flat disc of 3-nm thickness and 50-nm diameter. This condensate comprises one molecule of the plasmid DNA and about 2000 molecules of the fullerene derivative. Upon further addition of fullerene **5**, many single-molecule DNA condensates gather together to form the micrometer-sized object (Fig. 10c). Release of the DNA molecules from these large particles was achieved (Fig. 10d) by extractive removal of the fullerene with $CHCl_3$.

The formation of globular condensates composed of a single to a few DNA molecules is unique among other DNA-condensation studies based on lipid and dendrimeric molecules, which create much larger and less structurally defined DNA condensates.[80] Studies on a series of DNA-binding aminofullerenes recently showed that DNA condensation is required, but not sufficient for successful transfection.[60,61]

Fullerene-Binding Antibody

A fullerene binding protein has been elicited purposely.[81] Immunization of mice with a C_{60}-bovine thyroglobulin (>600 kDa) conjugate elicited the desired IgG isotype antibody. Crystal structure of Fab fragment of the antifullerene antibody revealed that a shape-complementary clustering of hydrophobic amino acids forms a fullerene binding site.[82] Modeling study shows participation of the induced fit mechanism in the binding process. Interestingly, a subpopulation of the antibodies also recognizes C_{70} and even carbon nanotubes.[83–85]

CONCLUSION

The development of the science of water-soluble fullerenes has been very rapid. The molecular design starting from the compounds **1** and **2** has become much more elaborate and the synthesis has become more sophisticated, leading recently to the synthesis of the most complex compounds in high yield and on a large scale. Practical applications to clinical and materials uses are therefore just ahead. The science of water-soluble fullerene stimulated the studies to dissolve also carbon nanotubes in water and organic solvent.[86,87] The high cost of fullerenes has been the biggest barrier toward practical applications of any fullerene derivative, and this problem is now resolved by a large-scale production achieved by a Japanese company.[88] The cost problem is still a big issue in the use of metal-containing fullerenes, but may be resolved in the not-too-distant future.

REFERENCES

1. Osawa, E. Superaromaticity. Kagaku **1970**, *25* (9), 854–863.
2. Kroto, H.W.; Heath, J.R.; O'Brien, S.C.; Curl, R.F.; Smalley, R.E. C_{60}-buckminsterfullerene. Nature **1985**, *318* (6042), 162–163.
3. Krätschmer, W.; Lamb, L.D.; Fostiropoulos, K.; Huffman, D.R. Solid C_{60}—A new form of carbon. Nature **1990**, *347* (6291), 354–358.
4. Howard, J.B.; McKinnon, J.T.; Makarovsky, Y.; Lafleur, A.L.; Johnson, M.E. Fullerenes C_{60} and C_{70} in flames. Nature **1991**, *352* (6331), 139–141.
5. Tokuyama, H.; Yamago, S.; Nakamura, E.; Shiraki, T.; Sugiura, Y. Photo-induced biochemical activity of fullerene carboxylic acid. J. Am. Chem. Soc. **1993**, *115* (17), 7918–7919.
6. Friedman, S.H.; DeCamp, D.L.; Sijbesma, R.P.; Srdanov, G.; Wudl, F.; Kenyon, G.L. Inhibition of the HIV-1 protease by fullerene derivatives: Model building studies and experimental verification. J. Am. Chem. Soc. **1993**, *115* (15), 6506–6509.
7. Sijbesma, R.; Srdanov, G.; Wudl, F.; Castoro, J.A.; Wilkins, C.; Friedman, S.H.; DeCamp, D.L.; Kenyon, G.L. Synthesis of a fullerene derivative for the inhibition of HIV enzymes. J. Am. Chem. Soc. **1993**, *115* (15), 6510–6512.

8. Schinazi, R.F.; Sijbesma, R.; Srdanov, G.; Hill, C.L.; Wudl, F. Synthesis and virucidal activity of a water-soluble, configurationally stable, derivatized C_{60} fullerene. Antimicrob. Agents Chemother. **1993**, *37* (8), 1707–1710.
9. Jensen, A.W.; Wilson, S.R.; Schuster, D.I. Biological applications of fullerenes: A review. Bioorg. Med. Chem. **1996**, *4* (6), 767–779.
10. Da Ros, T.; Prato, M. Medicinal chemistry with fullerenes and fullerene derivatives. Chem. Commun. **1999**, (8), 663–669.
11. Wilson, S.R. Nanomedicine: Fullerene and Carbon Nanotube Biology. In *Perspectives of Fullerene Nanotechnology*; Osawa, E., Ed.; Kluwer Academic: Dordrecht, 2002; 155–163.
12. Hamza, A.V.; Balooch, M. The chemisorption of C_{60} on Si(100)-(2 × 1). Chem. Phys. Lett. **1993**, *201* (5–6), 404–408.
13. Deguchi, S.; Alargova, R.G.; Tsuji, K. Stable dispersions of fullerenes, C_{60} and C_{70}, in water. Preparation and characterization. Langmuir **2001**, *17* (19), 6013–6017.
14. Guldi, D.M.; Prato, M. Excited-state properties of C_{60} fullerene derivatives. Acc. Chem. Res. **2000**, *33* (10), 695–703.
15. Martin, N.; Sánchez, L.; Illescas, B.; Pérez, I. C_{60}-based electroactive organofullerenes. Chem. Rev. **1998**, *98* (7), 2527–2547.
16. Hirsch, A. *Chemistry of the Fullerenes*; Thieme: Stuttgart, 1994.
17. Diederich, F.; Kessinger, R. Regio- and Stereoselective Multiple Functionalization of Fullerenes. In *Templated Organic Synthesis*; Diederich, F., Stang, P.J., Eds.; Wiley-VCH: Weinheim, 2000; 189–218.
18. Ruoff, R.S.; Tse, D.S.; Malhorta, R.; Lorents, D.C. Solubility of C_{60} in a variety of solvents. J. Phys. Chem. **1993**, *97*, 3379–3383.
19. Korobov, M.V.; Smith, A.L. Solubility of the Fullerenes. In *Fullerenes: Chemistry, Physics, and Technology*; Kadish, K.M., Ruoff, R.S., Eds.; Wiley: New York, 2000; 53–89.
20. Ruoff, R.S.; Malhotra, R.; Huestis, D.L.; Tse, D.S.; Lorents, D.C. Anomalous solubility behaviour of C_{60}. Nature **1993**, *362* (6416), 140–141.
21. Scrivens, W.A.; Tour, J.M. Potent solvents for C_{60} and their utility for the rapid acquisition of ^{13}C NMR data for fullerenes. J. Chem. Soc., Chem. Commun. **1993**, (15), 1207–1208.
22. Sivaraman, N.; Dhamodaran, R.; Kaliappan, I.; Srinivasan, T.G.; Rao, P.R.V.; Mathews, C.K. Solubility of C_{60} in organic solvents. J. Org. Chem. **1992**, *57* (22), 6077–6079.
23. Yamakoshi, Y.N.; Yagami, T.; Fukuhara, K.; Sueyoshi, S.; Miyata, N. Solubilization of fullerenes into water with poly(vinylpyrrolidone) applicable to biological tests. J. Chem. Soc., Chem. Commun. **1994**, (4), 517–518.
24. Diederich, F.; Gómez-López, M. Supramolecular fullerene chemistry. Chem. Soc. Rev. **1999**, *28* (5), 263–277.
25. Chiang, L.Y.; Upasani, R.B.; Swirczewski, J.W. Versatile nitronium chemistry for C_{60} fullerene functionalization. J. Am. Chem. Soc. **1992**, *114* (26), 10154–10157.
26. Prato, M.; Suzuki, T.; Foroudian, H.; Li, Q.; Khemani, K.; Wudl, F.; Leonetti, J.; Little, R.D.; White, T.; Rickborn, B.; Yamago, S.; Nakamura, E. [3+2] and [4+2] Cycloadditions of C_{60}. J. Am. Chem. Soc. **1993**, *115* (4), 1594–1595.
27. Irie, K.; Nakamura, Y.; Ohigashi, H.; Tokuyama, H.; Yamago, S.; Nakamura, E. Photocytotoxicity of water-soluble fullerene derivatives. Biosci. Biotechnol. Biochem. **1996**, *60* (8), 1359–1361.
28. Nakamura, E.; Tokuyama, H.; Yamago, S.; Shiraki, T.; Sugiura, Y. Biological activity of water-soluble fullerenes. Structural dependence of DNA cleavage, cytotoxicity and enzyme inhibitory activities. Bull. Soc. Chem. Jpn. **1996**, *69* (8), 2143–2151.
29. Yamakoshi, Y.; Sueyoshi, S.; Fukuhara, K.; Miyata, N.; Masumizu, T.; Kohno, M. $^{\bullet}OH$ and $O_2^{\bullet-}$ Generation in aqueous C_{60} and C_{60} solutions by photoirradiation: An EPR study. J. Am. Chem. Soc. **1998**, *120* (47), 12363–12364.
30. Bernstein, R.; Prat, F.; Foote, C. On the mechanism of DNA cleavage by fullerenes investigated in model systems: Electron transfer from guanosine and 8-oxo-guanosine derivatives to C_{60}. J. Am. Chem. Soc. **1999**, *121* (2), 464–465.
31. Nakamura, E.; Isobe, H.; Tomita, N.; Sawamura, M.; Jinno, S.; Okayama, H. Functionalized fullerene as a new artificial vector for transfection. Angew. Chem., Int. Ed. **2000**, *39* (23), 4254–4257.
32. Brettreich, M.; Hirsch, A. A highly water-soluble dendro[60]fullerene. Tetrahedron Lett. **1998**, *39* (18), 2731–2734.
33. Schuster, D.I.; Wilson, S.R.; Kirschner, A.N.; Schinazi, R.F.; Schlueter-Wirtz, S.; Tharnish, P.; Barnett, T.; Ermolieff, J.; Tang, J.; Brettreich, M.; Hirsch, A. Evaluation of the anti-HIV potency of a water-soluble dendrimeric fullerene. Proc. Electrochem. Soc. **2000**, *11* (9), 267–270.
34. Nelson, M.A.; Domann, F.E.; Bowden, G.T.; Hooser, S.B.; Fernando, Q.; Carter, D.E. Effects of acute and subchronic exposure of topically

applied fullerene extracts on the mouse skin. Toxicol. Ind. Health **1993**, *9* (4), 623–630.
35. Scrivens, W.A.; Tour, J.M.; Creek, K.E.; Pirisi, L. Synthesis of ^{14}C-labeled C_{60}, its suspension in water, and its uptake by human keratinocytes. J. Am. Chem. Soc. **1994**, *116* (19), 4517–4518.
36. Moussa, F.; Roux, S.; Pressac, M.; Génin, E.; Hadchouel, M.; Trivin, F.; Rassat, A.; Céolin, R.; Szwarc, H. In vivo reaction between [60]fullerene and vitamin A in mouse liver. New J. Chem. **1998**, *22* (9), 989–992.
37. Nakamura, E.; Yamago, S.; Ejiri, S.; Dorigo, A.E.; Morokuma, K. Reversible generation of trimethylenemethanes by mild thermolysis of dialkoxy methylenecyclopropanes. J. Am. Chem. Soc. **1991**, *113* (8), 3183–3184.
38. Yamago, S.; Tokuyama, H.; Nakamura, E.; Kikuchi, K.; Kananishi, S.; Sueki, K.; Nakahara, H.; Enomoto, S.; Ambe, F. In vivo biological behavior of a water-miscible fullerene: ^{14}C labeling, absorption, distribution, excretion and acute toxicity. Chem. Biol. **1995**, *2* (6), 385–389.
39. Nakamura, E.; Yamago, S. Thermal reactions of dipolar trimethylenemethane species. Acc. Chem. Res. **2002**, *35* (10), 867–877.
40. Yamago, S.; Nakamura, E. [3+2] Cycloaddition of trimethylenemethane and its synthetic equivalents. Org. React. **2002**, *61*, 1–215.
41. Dugan, L.L.; Turetsky, D.M.; Du, C.; Lobner, D.; Wheeler, M.; Almli, C.R.; Shen, C.K.-F.; Luh, T.-Y.; Choi, D.W.; Lin, T.-S. Carboxyfullerenes as neuroprotective agents. Proc. Natl. Acad. Sci. U. S. A. **1997**, *94* (17), 9434–9439.
42. Cagle, D.W.; Kennel, S.J.; Mirzadeh, S.; Alford, J.M.; Wilson, L.J. In vivo studies of fullerene-based materials using endohedral metallofullerene radiotracers. Proc. Natl. Acad. Sci. U. S. A. **1999**, *96* (9), 5182–5187.
43. Rajagopalan, P.; Wudl, F.; Schinazi, R.F.; Boudinot, F.D. Pharmacokinetics of a water-soluble fullerene in rats. Antimicrob. Agents Chemother. **1996**, *40* (10), 2262–2265.
44. Rancan, F.; Rosan, S.; Boehm, F.; Cantrell, A.; Brellreich, M.; Schoenberger, H.; Hirsch, A.; Moussa, F. Cytotoxicity and photocytotoxicity of a dendritic C_{60} mono-adduct and a malonic acid C_{60} tris-adduct on Jurkat cells. J. Photochem. Photobiol., B Biol. **2002**, *67* (3), 157–162.
45. Tzeng, S.-F.; Lee, J.-L.; Kuo, J.-S.; Yang, C.-S.; Murugan, P.; Tai, L.A.; Hwang, K.C. Effects of malonate C_{60} derivatives on activated microglia. Brain Res. **2002**, *940* (1–2), 61–68.
46. Chen, H.H.C.; Yu, C.; Ueng, T.H.; Chen, S.; Chen, B.J.; Huang, K.J.; Chiang, L.Y. Acute and subacute toxicity study of water-soluble polyalkylsulfonated C_{60} in rats. Toxicol. Pathol. **1998**, *26* (1), 143–151.
47. Buseck, P.R.; Tsipursky, S.J.; Hettich, R. Fullerenes from the geological environment. Science **1992**, *257* (5067), 215–217.
48. Chibante, L.P.F.; Heymann, D. On the geochemistry of fullerenes: Stability of C_{60} in ambient air and the role of ozone. Geochim. Cosmochim. Acta **1993**, *57* (8), 1879–1881.
49. Boutorine, A.S.; Tokuyama, H.; Takasugi, M.; Isobe, H.; Nakamura, E.; Hélène, C. Fullerene-oligonucleotide conjugates: Photo-induced sequence-specific DNA cleavage. Angew. Chem., Int. Ed. Engl. **1994**, *33* (23/24), 2462–2465.
50. An, Y.Z.; Chen, C.H.B.; Anderson, J.L.; Sigman, D.S.; Foote, C.S.; Rubin, Y. Sequence-specific modification of guanosine in DNA by a C_{60}-linked deoxyoligonucleotide: Evidence for a non-singlet oxygen mechanism. Tetrahedron **1996**, *52* (14), 5179–5189.
51. Da Ros, T.; Bergamin, M.; Vazquez, E.; Spalluto, G.; Baiti, B.; Moro, S.; Boutorine, A.; Prato, M. Synthesis and molecular modeling studies of fullerene-5,6,7-trimethoxyindole-oligonucleotide conjugates as possible probes for study of photochemical reactions in DNA triple helices. Eur. J. Org. Chem. **2002**, (3), 405–413.
52. Yamakoshi, Y.N.; Yagami, T.; Sueyoshi, S.; Miyata, N. Acridine adduct of [60]fullerene with enhanced DNA-cleaving activity. J. Org. Chem. **1996**, *61* (21), 7236–7237.
53. Tabata, Y.; Murakami, Y.; Ikada, Y. Photodynamic effect of polyethylene glycol-modified fullerene on tumor. Jpn. J. Cancer Res. **1997**, *88* (11), 1108–1116.
54. Tabata, Y.; Ishii, T.; Aoyama, T.; Oki, R.; Hirano, Y.; Ogawa, O.; Ikada, Y. Sonodynamic Effect of Polyethylene Glycol-Conjugated Fullerene on Tumor, Perspectives of Fullerene Nanotechnology. In *Perspectives of Fullerene Nanotechnology*; Osawa, E., Ed.; Kluwer Academic: Dordrecht, 2002; 185–196.
55. Chi, Y.; Canteenwala, T.; Chen, H.H.C.; Jeng, U.-S.; Lin, T.-L.; Chiang, L.Y. Free Radical Scavenging and Photodynamic Functions of Micelle-Like Hydrophilic Hexa (Sulfobutyl)Fullerene (FC4S). In *Perspectives of Fullerene Nanotechnology*; Osawa, E., Ed.; Kluwer Academic: Dordrecht, 2002; 165–183.
56. Kiselev, O.I.; Kozeletskaya, K.N.; Melenevskaya, E.Y.; Vinogradova, L.V.; Kever, E.E.; Klenin, S.I.; Zgonnik, V.N.; Dumpis, M.A.; Piotrovsky, L.B.

O

Antiviral activity of fullerene C_{60} with the poly(*N*-vinylpyrrolidone) complex. Mol. Cryst. Liq. Cryst. Sci. Technol., Sect. C Mol. Mater. **1998**, *11* (1–2), 121–124.

57. Wolff, J.A. *Gene Therapeutics: Methods and Applications of Direct Gene Transfer*; Birkhäuser: Boston, 1993.
58. Isobe, H.; Tomita, N.; Jinno, S.; Okayama, H.; Nakamura, E. Synthesis and transfection capability of multi-functionalized fullerene polyamine. Chem. Lett. **2001**, (12), 1214–1215.
59. Isobe, H.; Tokuyama, H.; Sawamura, M.; Nakamura, E. Synthetic and computational studies on symmetry-defined double cycloaddition of a new tris-annulating reagent to C_{60}. J. Org. Chem. **1997**, *62* (15), 5034–5041.
60. Isobe, H.; Tomita, N.; Nakamura, E. One-step multiple-addition of amine to [60]fullerene. Synthesis of tetra (amino)fullerene epoxide under photochemical aerobic conditions. Org. Lett. **2000**, *2* (23), 3663–3665.
61. Isobe, H.; Ohbayashi, A.; Sawamura, M.; Nakamura, E. A cage with fullerene end caps. J. Am. Chem. Soc. **2000**, *122* (11), 2669–2670.
62. Cassell, A.M.; Scrivens, W.A.; Tour, J.M. Assembly of DNA/fullerene hybrid materials. Angew. Chem., Int. Ed. **1998**, *37* (11), 1528–1531.
63. Cassell, A.M.; Asplund, C.L.; Tour, J.M. Self-assembling supramolecular nanostructures from a C_{60} derivative: Nanorods and vesicles. Angew. Chem., Int. Ed. **1999**, *38* (16), 2403–2405.
64. Chiang, L.Y.; Lu, F.-J.; Lin, J.-T. Free radical scavenging activity of water-soluble fullerenols. J. Chem. Soc., Chem. Commun. **1995**, (12), 1283–1284.
65. Kato, H.; Kanazawa, Y.; Okumura, M.; Taninaka, A.; Yokawa, T.; Shinohara, H. Lanthanoid endohedral metallofullerenols for MRI contrast agents. J. Am. Chem. Soc. **2003**, *125* (14), 4391–4397.
66. Mikawa, M.; Kato, H.; Okumura, M.; Narazaki, M.; Kanazawa, Y.; Miwa, N.; Shinohara, H. Paramagnetic water-soluble metallofullerenes having the highest relaxivity for MRI contrast agents. Bioconjug. Chem. **2001**, *12* (4), 510–514.
67. Bolskar, R.D.; Benedetto, A.F.; Husebo, L.O.; Price, R.E.; Jackson, E.F.; Wallace, S.; Wilson, L.J.; Alford, J.M. First soluble $M@C_{60}$ derivatives provide enhanced access to metallofullerenes and permit in vivo evaluation of $Gd@C_{60}[C(COOH)_2]_{10}$ as a MRI contrast agent. J. Am. Chem. Soc. **2003**, *125* (18), 5471–5478.
68. Mirkin, C.A.; Caldwell, W.B. Thin film, fullerene-based materials. Tetrahedron **1996**, *52* (14), 5113–5130.
69. Matsumoto, M.; Tachibana, H.; Azumi, R.; Tanaka, M.; Nakamura, T.; Yunome, G.; Abe, M.; Yamago, S.; Nakamura, E. Langmuir–Blodgett film of amphiphilic C_{60} carboxylic acid. Langmuir **1995**, *11* (2), 660–665.
70. Sano, M.; Oishi, K.; Ishi-i, T.; Shinkai, S. Vesicle formation and its fractal distribution by bola-amphiphilic [60]fullerene. Langmuir **2000**, *16* (8), 3773–3776.
71. Georgakilas, V.; Pellarini, F.; Prato, M.; Guldi, D.M.; Mlle-Franco, M.; Zerbetto, F. Supramolecular self-assembled fullerene nanostructures. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99* (8), 5075–5080.
72. Sawamura, M.; Iikura, H.; Nakamura, E. The first pentahapto fullerene metal complexes. J. Am. Chem. Soc. **1996**, *118* (50), 12850–12851.
73. Nakamura, E.; Sawamura, M. Chemistry of η^5-fullerene metal complexes. Pure Appl. Chem. **2001**, *73* (2), 355–359.
74. Sawamura, M.; Iikura, H.; Hirai, A.; Nakamura, E. Synthesis of π-indenyl-type fullerene ligand and its metal complexes via quantitative trisarylation of C_{70}. J. Am. Chem. Soc. **1998**, *120* (32), 8285–8286.
75. Sawamura, M.; Nagahama, N.; Toganoh, M.; Hackler, U.E.; Isobe, H.; Nakamura, E.; Zhou, S.-Q.; Chu, B. Pentaorgano[60]fullerene $R_5C_{60}^-$. A water soluble hydrocarbon anion. Chem. Lett. **2000**, (9), 1098–1099.
76. Zhou, S.; Burger, C.; Chu, B.: Sawamura, M.; Nagahama, N.; Toganoh, M.; Hackler, U.E.; Isobe, H.; Nakamura, E. Spherical bilayer vesicles of fullerene based surfactants in water: A laser light scattering study. Science **2001**, *291* (5510), 1944–1947.
77. Noguchi, H.; Takasu, M. Self-assembly of amphiphiles into vesicles: A Brownian dynamics simulation. Phys. Rev., E **2001**, *64* (5), 0419131–0419137.
78. Sawamura, M.; Kawai, K.; Matsuo, Y.; Kanie, K.; Kato, T.; Nakamura, E. Stacking of conical mesogens with a fullerene apex into polar columns in crystals and liquid crystals. Nature **2002**, *419* (6908), 702–705.
79. Isobe, H.; Sugiyama, S.; Fukui, K.-I.; Iwasawa, Y.; Nakamura, E. Atomic force microscope studies on condensation of plasmid DNA with functionalized fullerene. Angew. Chem., Int. Ed. **2001**, *40* (18), 3364–3367.
80. Kawaura, C.; Noguchi, A.; Furuno, T.; Nakanishi, M. Atomic force microscopy for studying gene

transfection mediated by cationic liposomes with a cationic cholesterol derivative. FEBS Lett. **1998**, *421* (1), 69–72.
81. Chen, B.-X.; Wilson, S.R.; Das, N.; Coughlin, D.J.; Erlanger, B.F. Antigenicity of fullerenes: Antibodies specific for fullerenes and their characteristics. Proc. Natl. Acad. Sci. U. S. A. **1998**, *95* (18), 10809–10813.
82. Braden, B.C.; Goldbaum, F.A.; Chen, B.-X.; Kirschner, A.N.; Wilson, S.R.; Erlanger, B.F. X-ray crystal structure of an anti-buckminsterfullerene antibody Fab fragment: Biomolecular recognition of C_{60}. Proc. Natl. Acad. Sci. U. S. A. **2000**, *97* (22), 12193–12197.
83. Erlanger, B.F.; Brus, L.; Sheetz, M.P. Antibodies specific to fullerene nanotubes for detection and immobilization of nanotubes. PCT Int. Appl. **2003**, 155.
84. Erlanger, B.F.; Chen, B.-X.; Zhu, M.; Brus, L. Binding of an anti-fullerene IgG monoclonal antibody to single wall carbon nanotubes. Nano Lett. **2001**, *1* (9), 465–467.
85. Pantarotto, D.; Partidos, C.D.; Graff, R.; Hoebeke, J.; Briand, J.-P.; Prato, M.; Bianco, A. Synthesis, structural characterization, and immunological properties of carbon nanotubes functionalized with peptides. J. Am. Chem. Soc. **2003**, *125* (29), 6160–6164.
86. Bahr, J.L.; Tour, J.M. Covalent chemistry of single-wall carbon nanotube. J. Mater. Chem. **2002**, *12* (7), 1952–1958.
87. Hirsch, A. Functionalization of single-walled carbon nanotubes. Angew. Chem., Int. Ed. **2002**, *41* (11), 1853–1859.
88. Frontier Carbon Corporation. http://www.f-carbon.com.

Oxide Nanoparticles: Electrochemical Performance

Dominique Larcher
Jean-Marie Tarascon
Université de Picardie—Jules Verne, Amiens, France

INTRODUCTION

Although intensively studied in a variety of fields such as catalysis, magnetism, and optics, oxide nanoparticles have just been modestly considered within the field of energy storage until these last few years that have witnessed intense developments, mainly regarding Li-ion battery technology as devoted herein.

For the last 20 years, it was a common and well-accepted belief that highly divided materials could not be suitable for extended reversible redox reaction with metallic lithium. This seems quite astonishing because most primary and secondary lithium electrochemical cell devices are based on redox reactions involving interfacial reactions between a liquid organic-based electrolyte and a solid electrode material. In fact, this idea was rooted on the belief that electrochemically driven irreversible decomposition of the organic-based electrolyte occurs at the surface of the active particles while the cell is cycled, together with possible dissolution of the solids. Thus, the higher the surface of contact between the particles and the electrolyte, the higher would be the extent of this capacity loss, hence the requirement of large particles with low surface area. Along that line, the BET surface areas of powders presently used in commercial Li-ion cells do not exceed 2 m^2/g. Recently, these prevailing ideas were seriously contradicted by several findings linked to the recent interest in nanopowders, as we will discuss in this paper. Owing to the present staggering trend toward nano objects in various research fields, too often driven by funding opportunities rather than true science, much confusion has surged about the real meaning of "nano." Undoubtedly, the world of the so-called nanosciences is still waiting for a concrete and universal definition or at least an accurate, related size scale. Thus, we talk about nanopowders, nanostructures/nanocomposites, nanotextures, and nanoarchitectured electrodes by reference to materials having a single component, two or more components intimately mixed at the nanometric scale, a porous electrode having pores and components in the nanometric scale, and finally a well-designed (either by template deposition or by lithography) two-component system at the nanometric scale. Basically, we are prone to think that a given object, observed for a given property with a specific characterization mean, enters the "nano-world" as soon as its size reaches a value below which the studied property starts to drastically differ from that of bulk. For instance, the significant decrease in Au particles' melting point as their particle size becomes smaller than 50 Å perfectly illustrates this point,[1] and for such a phenomenon 50 Å will be the threshold value delimiting the nano/macro worlds. It is therefore unfortunate that we do not have such an equivalent to Plank's constant that neatly separates classical mechanics from quantum mechanics.

As we march from bulk materials toward systems with small particles, new electrochemical effects are recently observed and these require the formulation of a new theoretical foundation on which to base our interpretation and understanding of the experimental data. A theoretical approach was recently undertaken and the effect of surface, particle size toward the evolution of electrochemical reactions considered. A profound effect on the chemical potential–composition curves and energy sites was demonstrated.[2] Moreover, a decrease in domain size results in a dramatic increase in grain boundaries[3] that were shown, by acting as space charge regions, to be beneficial to ionic conductivity. However, the particle size is not the only parameter that can account for the modification of properties. Indeed, the size of the crystallized reacting domains (crystallites), the specific surface area of the powder, the porosity, and the possible confinement of the matter can have a drastic influence as well.

It is surprising that the effect of the size of the reacting domains on the electrochemical properties remains almost uninvestigated, while most of the scientific fields are now, as a whole, turning toward the nanomaterials. Given the implications of the energy storage in our evolving modern society (communication, biocompatible devices), we definitively have to consider this rich domain of opportunities that lies beyond the use of nanomaterials in this field, as illustrated by the following examples.

LITHIUM REACTIVITY THROUGH INSERTION PROCESSES

Although the principle of a Li-based secondary cell dates back to the 70s, we had to wait until the 90s for the first

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013825

commercialization of reliable and safe rechargeable batteries. This impressive delay originates from numerous scientific and technological problems whose step-by-step resolution lasted decades and required sustained and crossed endeavor from physicists, chemists, and theorists. Aside from the formulation of stable electrolytes within a wide electrochemical potential window, the search for suitable electrode materials is still a major issue. The main difficulty encountered along this quest for the "ideal" insertion material is nested in the numerous criteria they have to meet all in one. They have to show high insertion capacity without undergoing drastic structural changes, low toxicity and low cost, highly separated redox potential between the cathode and anode, enhanced thermal/chemical stability toward the electrolyte, high electronic/ionic conductivities for power rate capabilities, and, finally, easy synthesis with scale-up possibilities.

Such "ideal" materials do not exist and to date only very few compounds were isolated as coming close to these requirements: ternary $LiCoO_2$, $LiNiO_2$, $LiMn_2O_4$ oxides for the cathode,[4,5] and carbonaceous materials for the anode. To bypass the identified drawbacks of these materials, present studies dealing with anionic and cationic substitutions,[6] surface treatments,[7] or the making of composite materials are exploring new compromises rather than innovative and creative approaches, with the exception of a few that are exploiting lithium chemical reactivity at the nanometric scale.

Below, we give some examples showing how nanomaterials, either as electrodes or as part of the electrolyte composition, could greatly benefit the field of rechargeable Li batteries.

$LiMn_2O_4$spinel oxide is the first example selected to illustrate how Li-driven phase transitions associated to drastic electrode volume changes and leading to poor electrochemical performances can turn to our advantage when using small reacting domains. It has long been well known that Li insertion in this spinel framework results in a cooperative Jahn–Teller distortion together with the formation of a tetragonal phase showing a volume expansion of about 16% as compared to its cubic precursor with the overall result being very poor retention capacity upon cycling.[4,5] In contrast, when very small internal crystallites are formed by intense mechanical milling performed on crystallized powders or by precipitation techniques, this reaction appears to be very reversible.[8,9] This unexpected behavior was attributed to an isotropic and benign expansion of the small domains as compared to the highly anisotropic and strain-inducing expansion of large ones, when the material is made by the "bake and shake" method. Because of the plasticity of the ball-milled powders, the electrode maintains its integrity upon insertion–removal of lithium ions with limited resulting loss of electrical contact.

By acting on the particle size, aside from the possibility of modifying the side effects associated to an intercalation phenomenon, one can modify the insertion mechanism by itself. This has been recently achieved for several compounds by acting on the size of the reacting domains as demonstrated through a few studies on carbonaceous materials.

Carbonaceous materials, owing to their both faradic and capacitive properties, are now widely implemented in Li-ion batteries and supercondensators. Their electrochemical properties are highly dependent on their texture, composition, and synthetic history. Although capacitor technology has long recognized the benefits of high-surface-area materials, it was quite recently that Li insertion in disordered and highly divided carbons turned out to be a promising approach. Such disordered/divided carbonaceous materials were, for instance, elaborated by direct preparation through pyrolysis of organic precursors or by postgrinding treatment under various conditions (atmosphere, grinding time, miller type) of crystallized carbons previously prepared at high temperatures.[10] Following the latter approach after an early appraisal of various ball-milling key parameters, carbonaceous materials having reversible specific electrochemical capacity that can reach 720 mA hr/g, which is about twice that of graphite, and low irreversible loss in the first cycle (200 mA hr/g) were fabricated. More importantly, such performance gains were not limited to one specific carbon precursor because ball milling, once the experimental parameters optimized, was shown to lead to identical texture whatever the carbonaceous precursor, stressing the universal effect of this mechanical treatment.[11] Several models have been proposed to explain this behavior, but

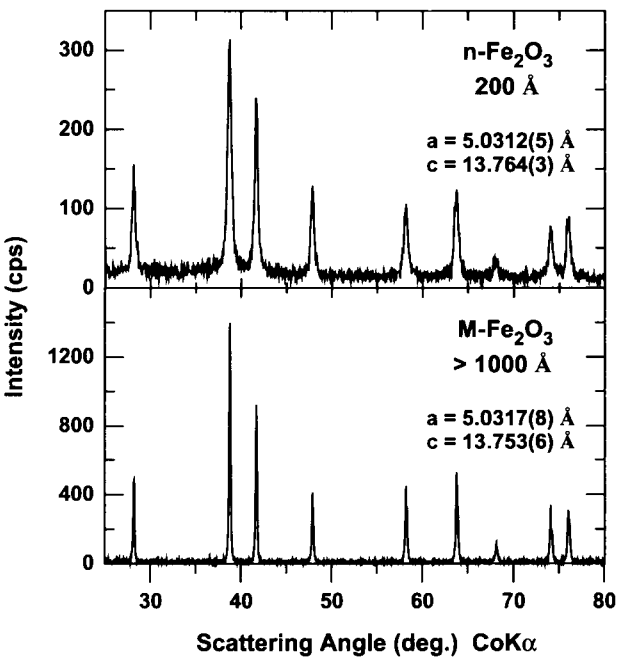


Fig. 1 X-ray diffraction patterns for n-Fe_2O_3 and M-Fe_2O_3.

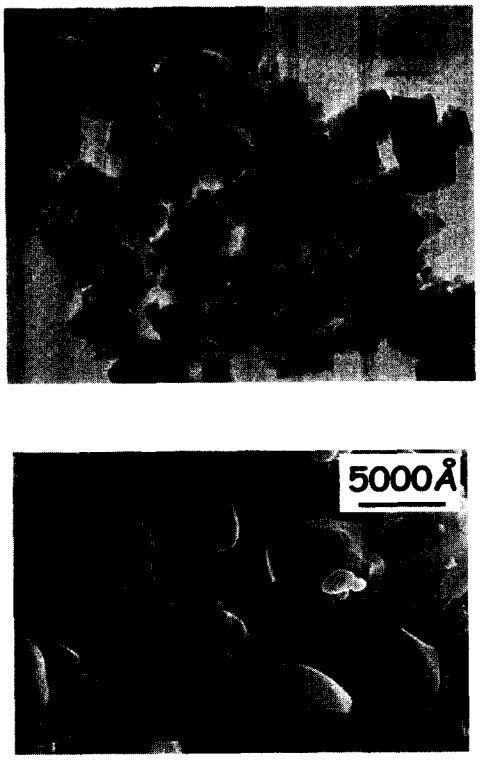


Fig. 2 TEM and SEM pictures of nanosized ($n\text{-}Fe_2O_3$) and micron-sized ($M\text{-}Fe_2O_3$) samples.

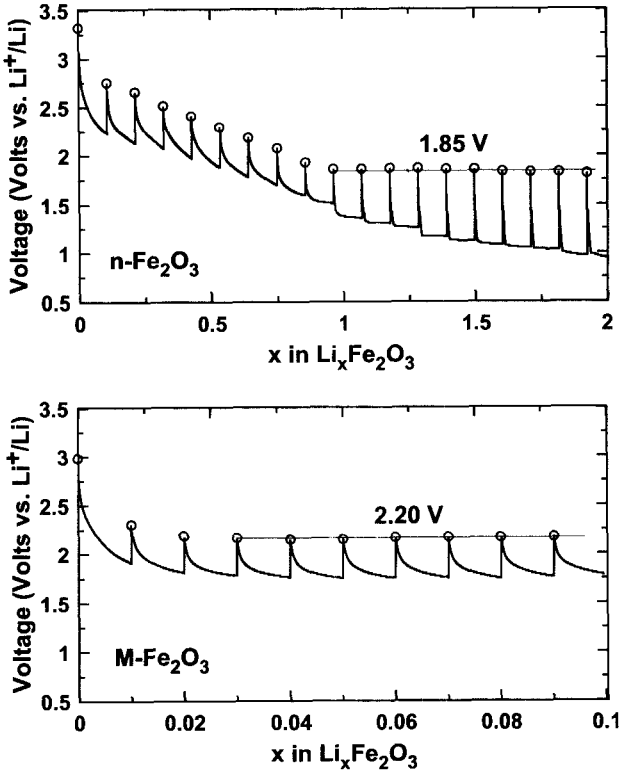


Fig. 3 Voltage–composition curves for $n\text{-}Fe_2O_3$ and $M\text{-}Fe_2O_3$ in Potentiostatic Intermittent Titration Technique mode. Open circles indicate the quasi-equilibrium OCVs.

despite this controversy, there is no doubt the origin of this increase in capacity is nested in the extremely small domains (30 Å) created by the milling.

Hematite $\alpha\text{-}Fe_2O_3$[12,13] and $M_3O_4(M = Fe, Co, Mn)$[15] will be used as pedagogical examples to illustrate the effect of particle size on the electrochemical reaction path itself upon reaction with metallic lithium.

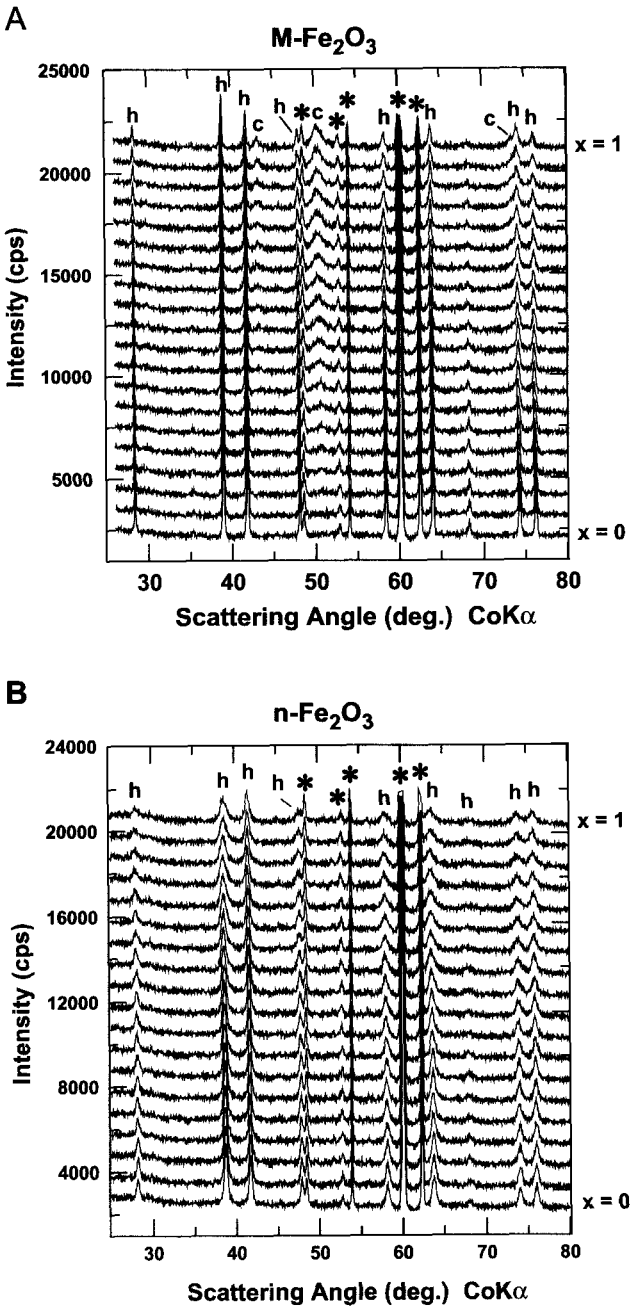


Fig. 4 In situ X-ray diffraction patterns collected on (A) Li/$M\text{-}Fe_2O_3$ and (B) Li/$n\text{-}Fe_2O_3$ cells discharged at C/5 rate. The "h" and "c" labels indicate the hexagonal $\alpha\text{-}Fe_2O_3$ and the cubic $Li_2Fe_2O_3$ phases, respectively. Asterisks indicate the peaks due to apparatus (Be, BeO).

The reaction of lithium with hematite α-Fe_2O_3 entails a very limited solid solution domain up to a lithium critical concentration of about 0.03 lithium per formula unit, and then an irreversible transition leading to a $Li_2Fe_2O_3$ rock-salt-type phase. It is surprising that when using very small hematite particles, this reactivity scheme appears to be drastically different. More specifically, two batches of hematite were compared: micron-sized (M-Fe_2O_3, 2 m^2/g, 0.5 μm mean particle size, commercial) and nanosized (n-Fe_2O_3, 60 m^2/g, 200 Å mean particle size, homemade) samples (Figs. 1 and 2). Quasi-equilibrium voltage evolutions along the reaction with lithium [open circuit voltage (OCV) measurements] confirm that the reduction of M-Fe_2O_3 proceeds through a biphasic process (2.20 V vs. Li^+/Li) after a very small solid solution domain. In contrast, a much longer single-phase solid solution domain (up to about one Li) is observed for n-Fe_2O_3, followed by a biphasic domain characterized by an OCV value of 1.85 V vs. $Li^+/Li°$ (Fig. 3). In situ X-ray diffraction experiments (Fig. 4) further confirm these differences in reactivity when the particle size is modified. For a maximum lithium amount of one per formula unit ($x=1$), M-Fe_2O_3 irreversibly transforms into a cubic rock-salt-type $Li_2Fe_2O_3$ phase, although the n-Fe_2O_3 Bragg reflections progressively and reversibly shift without any evidence for extra phase. The cell volume expansion linked to this topotactic insertion is evaluated to be 1% from α-Fe_2O_3 to α-$Li_1Fe_2O_3$. As expected for such benign cationic insertion associated with small volume changes, the nanometric material exhibits higher reversible capacity and improved cycling behavior when compared to the M-Fe_2O_3 batch (Fig. 5). Aside from these electrochemical aspects and the possibility of forming the α-$Li_1Fe_2O_3$ phase, recent Mössbauer and EXAFS measurements indicate that the intimate reaction path beyond $x=1$ is also dependent on the particle size. For instance, we could unambiguously demonstrate a disproportionation ($3Fe^{2+} \rightarrow 2Fe^{3+} + Fe°$) mechanism along the n-$Fe_2O_3$ reduction beyond $x=1$, whereas it is not observed for M-Fe_2O_3.[13]

Finally, the overall alternative paths observed depending on the domains size could be summarized as follows:

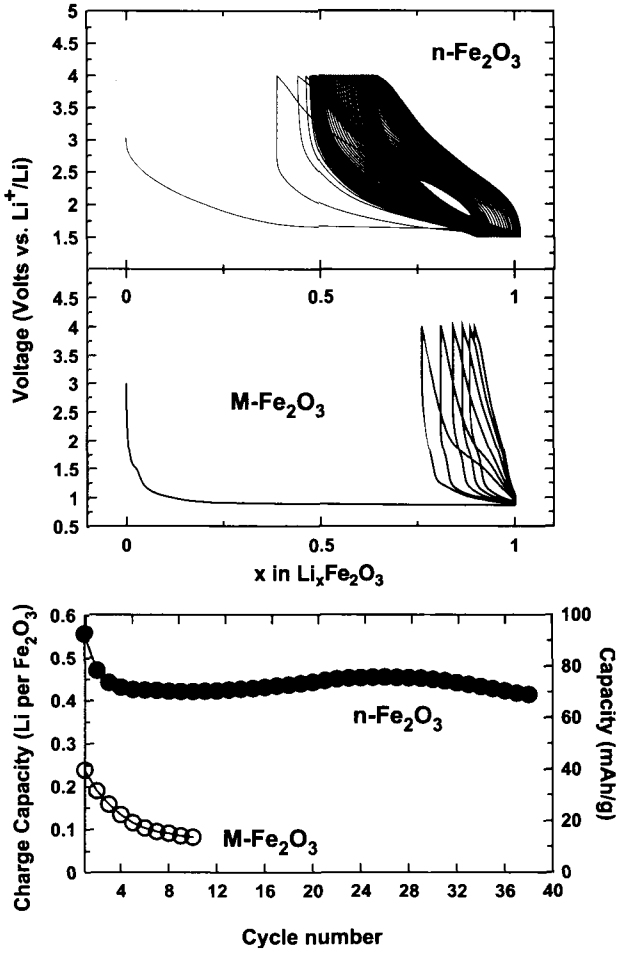
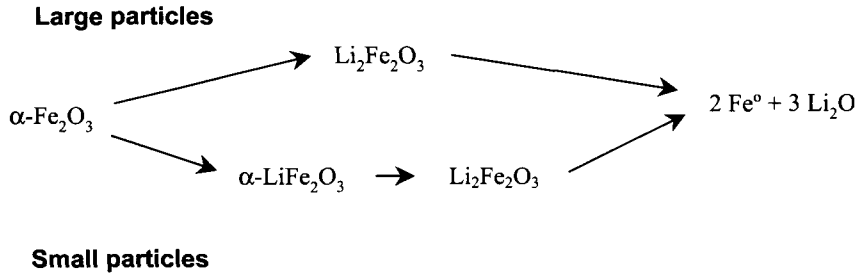
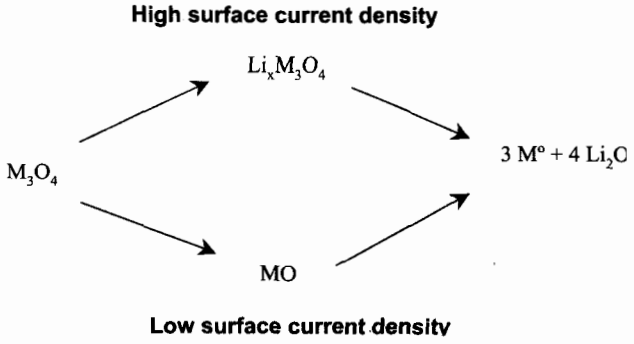


Fig. 5 Voltage–composition curves and capacity evolution upon cycling for Li/n-Fe_2O_3 and Li/M-Fe_2O_3 cells (1 Li/5 hr, 25°C).

Similar extension of the solid solution domain by using small particles/crystallites was also recently demonstrated for compact structures such as anatase TiO_2,[14] and we believe that numerous host materials disregarded for reversible lithium intercalation behave in a similar way. Thus, a completely new area is ahead of us in the search for new valuable insertion materials.



Besides hematite, numerous other materials with empty available sites are reported to irreversibly react with lithium. Among them, 3d-metal spinel materials M_3O_4 (M=Co, Mn, Fe) are also known to transform into lithiated rock-salt type $Li_xM_3O_4$ structures. This is illustrated in Fig. 6. Low-surface-area Co_3O_4 (3 m^2/g) reacts with metallic lithium at constant C/12 rate (one Li in 12 hr) through a two-phase Li insertion process leading to a cubic rock-salt-type phase with the approximate composition $Li_{\sim 1.5}Co_3O_4$ (Fig. 6A). In contrast, high-surface-area Co_3O_4 materials (66 m^2/g) behave differently because they transform into rock-salt α-CoO (Fig. 6B). The same CoO formation was observed when (3 m^2/g) Co_3O_4 powders were slowly reduced (C/100) or reacted with lithium in an intermittent mode consisting of short discharge steps followed by 2-hr relaxation periods. Note that similar current-density-driven observations were also made on Mn_3O_4 materials of different surface area. Such behavior fully contrasts with that of Fe_3O_4, as $Li_xFe_3O_4$ phases are always formed whatever the reduction rate or the texture of the starting Fe_3O_4 powders, suggesting an effect of the crystallographic nature of the spinel (normal for Co_3O_4 as compared to inverse for Fe_3O_4). For Co_3O_4 and Mn_3O_4, the key parameter governing the reduction reaction is therefore not the particle size but the applied current density (current/specific surface area ratio). Therefore, through these examples we demonstrate that by varying either the cycling rate or the specific surface area of the active material one can control the Li reacting path as summarized below.



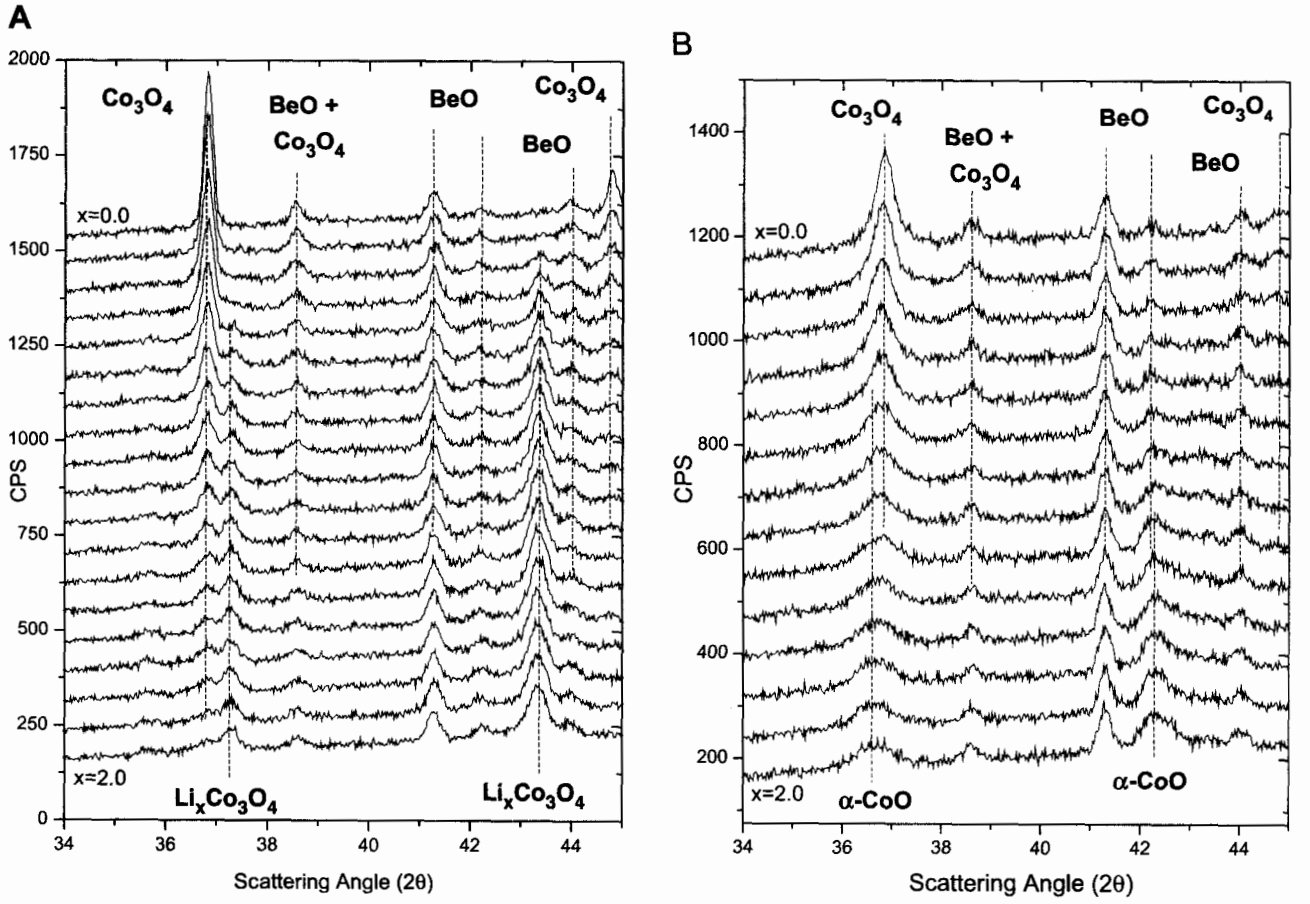


Fig. 6 In situ X-ray diffraction patterns collected during the electrochemical reduction of Low Surface Area-Co_3O_4 (A) and High Surface Area-Co_3O_4 (B) at constant current (1 Li/12 hr) at 25°C. The top patterns correspond to the initial Co_3O_4 and the bottom ones to x=2.0 Li per Co_3O_4 (CuK_α).

LITHIUM REACTIVITY THROUGH CONVERSION PROCESSES

Because the division state and texture of oxides used as electrode materials can drastically alter the initial intercalation process, one wonders whether the full reduction of 3d-metal oxides into metal can be affected as well. Although the size of the metallic clusters so formed are not significantly dependent on the initial oxide particle size, it was astonishing to observe, as demonstrated for various oxides such Cu_2O, CuO, CoO, NiO, MnO, FeO, Fe_2O_3, Fe_3O_4, and Co_3O_4,[16–18] a very high reversibility of these oxides toward Li according to the following reaction:

$$MO_x + 2xLi \Leftrightarrow M^\circ + xLi_2O$$

The reversibility of these conversion reactions was attributed to the in situ formation of nanometric metallic clusters embedded in a lithia matrix with, therefore, a very high metal/Li_2O surface of contact. Upon reoxidation of these nanocomposites, very small clusters of oxides are formed back. This is shown in Figs. 7–9 for CoO/Li cells. However, the formation of such nanocomposites is not the only key parameter required for high reversibility. Indeed, when ex situ formed highly divided oxides are cycled vs. lithium, the overall cell capacity drops very quickly. This is linked to the growing of an organic gel at the surface of the particles (resulting from the decomposition/reduction of the electrolyte) that tends to electrically isolate the active particles from the current collector. The criteria for good long-term cycling behavior are twofold: 1) low surface area of the initial oxide particles and 2) in situ formed nanodomains insuring high internal contact. Note that this need of in situ nanostructure was already demonstrated in aqueous electrochemical storage using nickel hydroxide as active material.[19] These reversible conversion reactions were not only limited to oxides but were found to occur for various classes of compounds such as nitrides, fluorides, and sulfides with formation of metal/Li_3N, metal/LiF, and metal/Li_2S nanocomposites, respectively.[20–24] Thanks to the universality of these reactions, to their appealing electrochemical performances, to their intrinsic simplicity from a thermodynamic point of view, and to the large panel of materials available, it clearly appears that they represent a main opportunity for future development in the field of electrode materials. Because they involve two to three electrons per 3d metal as compared to 1, and high voltages as well (3 V for CoF_3), such nanodriven reversible conversion reactions, owing to their staggering capacity gains over classical insertion reactions, could revolutionize the way the world stores and uses power.

ALLOYING REACTIONS

Lithium-alloying reaction with metal such as Al, Sn, Si, Ge, Cd, Pb, and Sb is one of the most attractive alternatives to intercalation reactions.[25–29] About 25 metals or semimetals are known to form alloys with lithium and most of their combinations are active toward lithium.[30–33] These reactions offer very high capacities (Si: 4200 mA hr/g, Ge: 1600 mA hr/g, Al and Sn: 990 mA hr/g) at relatively low voltages (<1 V vs. Li^+/Li°) but their cycling behavior is very poor. This detrimental characteristic is attributed to the pulverization of the material, and hence the loss of electrical contact, due to the large volume expansion of the active domains and high strain-inducing phase transitions together with the Li incorporation. Although these structural changes are intrinsic to the reactions, several attempts have been made to limit their side effects on the electrode integrity. Among them, the active/inactive nanocomposite concept represents one

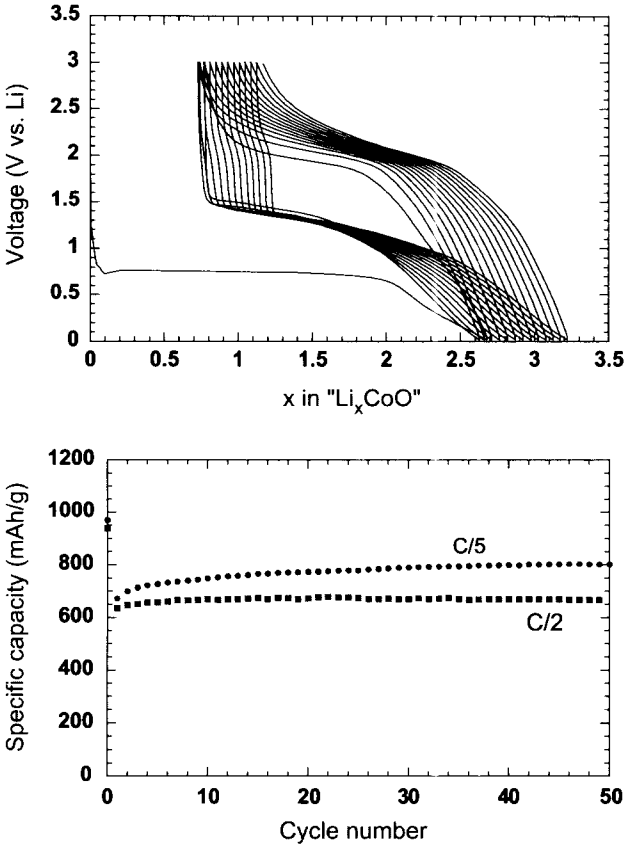


Fig. 7 Composition–voltage curve for a CoO/Li cell (top) and evolution of the capacity as a function of the cycle number (bottom). (*View this art in color at www.dekker.com.*)

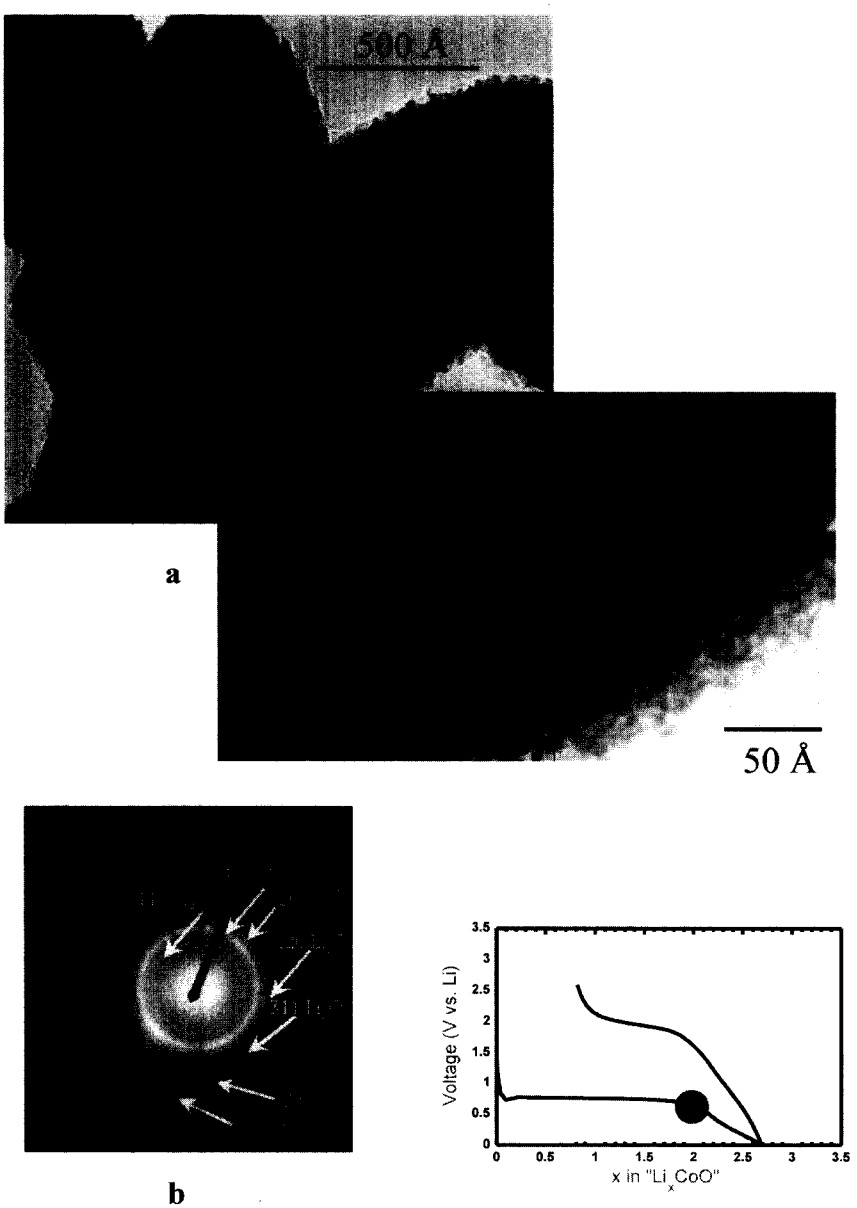


Fig. 8 HRTEM images (a) and corresponding selected-area electron diffraction (SAED) diagram of a CoO electrode after reaction with 2 Li per CoO. (*View this art in color at www.dekker.com.*)

of the most promising exploratory routes in the field. Such composites consist of two intimately mixed materials, one reacting with lithium while the other acts as an inactive confining buffer. In addition, the use of small-sized metallic clusters as Li hosts considerably suppress the associated strains and therefore improve their textural response to the alloying reaction.[34,35] Several recent examples can illustrate this concept through different systems such as Sn–O based glasses,[36–39] Si–C,[40–42] Sn–Fe–C[43–45] and Sn–Mn–C.[46,47] The best experimental proof of the beneficial aspect of the buffer matrix arises from the ability to obtain several hundreds of cycles on a composite made by precipitating Sn metal at the grain boundaries of an electrochemically inactive $SnFe_3C$ matrix. However, the retention capacity was improved at the expense of a decrease in the material electrochemical capacity. Undoubtedly, alloy performance can also benefit from nanostructuration, but further developments are

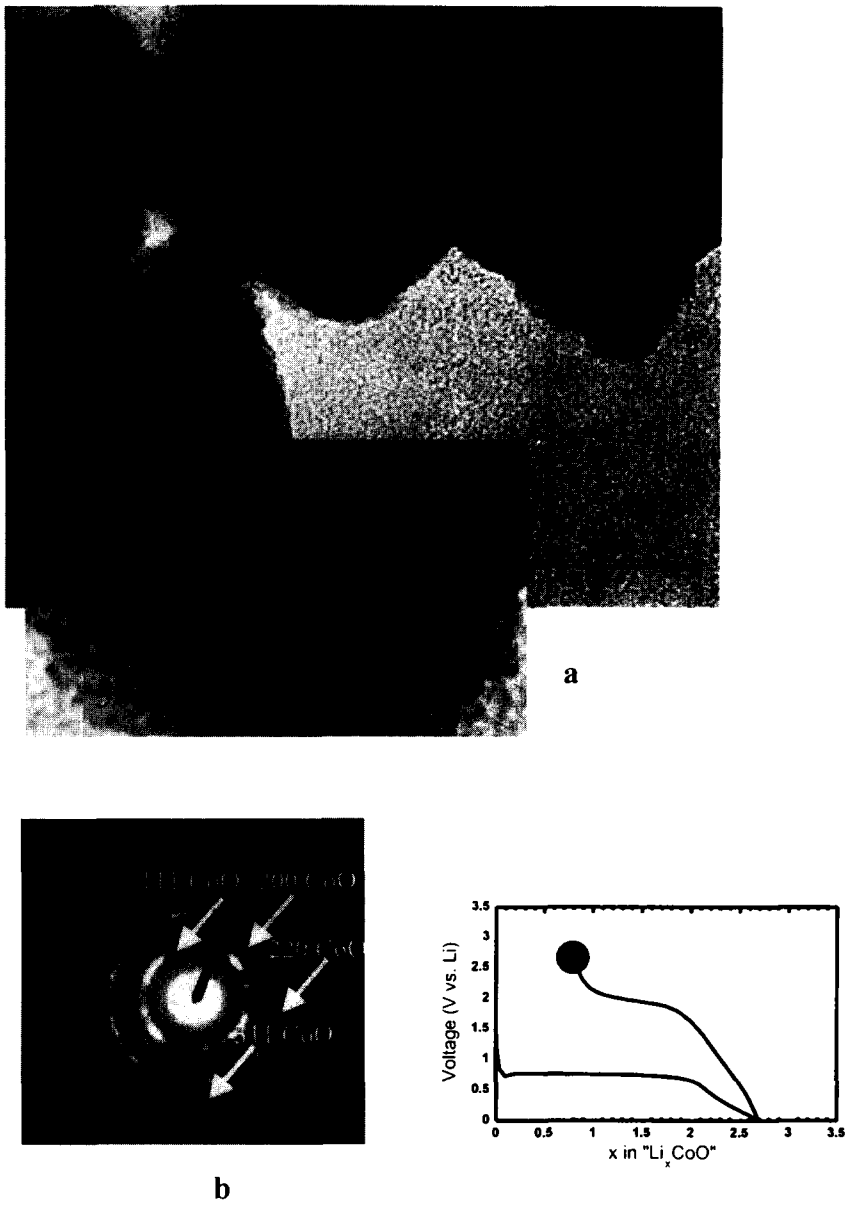


Fig. 9 HRTEM images (a) and corresponding SAED diagram of a CoO electrode after one discharge/charge cycle.

needed in the search for elaboration techniques to develop tailor-made nanostructure alloys.

ELECTROLYTE

Small particles are not only involved in electrode redox mechanisms but can also be of great implications for the electrolyte conduction properties. For instance, the addition of nanometric inorganic fillers (TiO_2, SiO_2, Al_2O_3) into various types of dry or hybrid polymer systems has proved effective in improving their ionic conduction properties to a large extent without affecting their interfacial and mechanical stabilities, in the whole temperature range.[48–50] This enhancement of the conductivity was explained both in chemical terms through the introduction of a Lewis acid–base concept that considers the effect of fillers in ion–ion and ion–polymer interactions in polymeric electrolytes and in theoretical terms through a modification of diffusion and transport coefficient in space charge regions created at the nanoceramic/polymer interfaces. Present trends are devoted toward searching for new nanostructured polymers loaded with functionalized ceramic nanomaterials (e.g., nanoceramics having acid or

basic surface groups). Such an approach was shown to be quite beneficial in developing highly performing protonic membranes for fuel cell applications.

CONCLUSION

It is now evident that the entire energy storage field is moving toward nanostructures and nanoparticles. So far, only the capacitor technology was apparently concerned by this aspect and it is probably not purely coincidental that recent works have demonstrated the feasibility of hybrid capacitive/faradic lithium cells.[51] Through this simple example, it is clear that a completely new area of opportunities is now open within the field of energy storage, both at the fundamental and applied levels. As we march from bulk materials toward systems with small particles, new electrochemical phenomena are observed and these require the formulation of a new theoretical foundation on which to base our interpretation and understanding of the experimental data. Regarding the applications, several foreseen benefits associated to nanomaterial-based Li batteries are the following: 1) an easier accommodation of structural strains for better calendar life, 2) a shorter diffusion path for enhanced electrode power capabilities, 3) an increase of the surface vs. core volume for larger double-layer capacitance contribution, and 4) an enhanced solid state reactivity to bypass kinetics limitations so that novel reactions showing staggering electrochemical capacity gains and involving low-cost materials are becoming feasible. Such cumulated effects should translate into tremendous gain costs while preserving safety. However, although quite exciting, enthusiasm should not prevail over reality, and a few critical problems that will require cooperative endeavor and cross-sharing of knowledge accumulated for the last decades in fields such as catalysis and others remain to be solved.

ACKNOWLEDGMENTS

The authors would like to give special thanks to L. Dupont, P. Poizot, S. Grugeon, S. Laruelle, E. Baudrin, G. Sudant (LRCS), and Y. Chabre (UJF) for their contributions to some of the data herein presented.

REFERENCES

1. Buffat, B.; Borel, J.P. Size effect on the melting temperature of gold particles. Phys. Rev., A **1976**, *13*, 2287–2292.
2. Obrovac, M.N.; Dahn, J.R. Implications of finite-size and surface effects on nanosize intercalation materials. Phys. Rev., B **2000**, *61* (10), 6713–6719.
3. Van Hardeveld, R.; Hartog, F. The statistics of surface atoms and surface sites on metal crystals. Surf. Sci. **1969**, *15*, 189–230.
4. Thackeray, M.M.; de Kock, A.; Rossouw, M.H.; Liles, D.; Bittihn, R.; Hoge, D. Spinel electrodes from the Li–Mn–O system for rechargeable lithium battery applications. J. Electrochem. Soc. **1992**, *139*, 363–366.
5. Thackeray, M.M. Manganese oxides for lithium batteries. Prog. Solid State Chem. **1997**, *25*, 1–71.
6. Amatucci, G.G.; Pereira, N.; Zheng, T.; Plitz, I.; Tarascon, J.M. Enhancement of the electrochemical properties of $Li_1Mn_2O_4$ through chemical substitution. J. Power Sources **1999**, *81–82*, 39–43.
7. Amatucci, G.G.; Blyr, A.; Sigala, C.; Alfonse, P.; Tarascon, J.M. Surface treatments of $Li_{1+x}Mn_{2-x}O_4$ spinels for improved elevated temperature performance. Solid State Ionics **1997**, *104* (1–2), 13–25.
8. Kim, J.; Manthiram, A. Nanocomposite manganese oxides for rechargeable lithium batteries. Electrochem. Solid-State Lett. **1998**, *1* (5), 207–209.
9. Kang, S.-H.; Goodenough, J.B.; Rabenberg, L.K. Nanocrystalline lithium manganese oxide spinel cathode for rechargeable lithium batteries. Electrochem. Solid-State Lett. **2001**, *4* (5), A49–A51.
10. Disma, F.; Aymard, L.; Dupont, L.; Tarascon, J.M. Effect of mechanical grinding on the lithium intercalation process in graphites and soft carbons. J. Electrochem. Soc. **1996**, *143* (12), 3959.
11. Salver-Disma, F.; Tarascon, J.M.; Clinard, C.; Rouzaud, J.N. Transmission electron microscopy studies on carbon materials prepared by mechanical milling. Carbon **1999**, *37*, 1941–1959.
12. Larcher, D.; Masquelier, C.; Bonnin, D.; Chabre, Y.; Masson, V.; Leriche, J.-B.; Tarascon, J.-M. Effect of particle size on lithium intercalation into α-Fe_2O_3. J. Electrochem. Soc. **2003**, *150* (1), A133–A139.
13. Larcher, D.; Bonnin, D.; Cortes, R.; Rivals, I.; Personnaz, L.; Tarascon, J.-M. Combined XRD, EXAFS and Mössbauer studies of the reduction by lithium of α-Fe_2O_3 with various particle size. J. Electrochem. Soc., *in press*.
14. Sudant, G.; Baudrin, E.; Larcher, D., Tarascon, J.-M. Electrochemical study of porous titanium oxide, submitted to *Chemistry of Materials*.
15. Larcher, D.; Sudant, G.; Leriche, J.-B.; Chabre, Y.; Tarascon, J.-M. The electrochemical reduction of Co_3O_4 in a lithium cell. J. Electrochem. Soc. **2002**, *149* (3), A234–A241.
16. Poizot, P.; Laruelle, S.; Grugeon, S.; Dupont, L.; Tarascon, J.M. Nano-sized transition-metal oxides

as negative-electrode materials for lithium-ion batteries. Nature **2000**, *407*, 496–499.
17. Badway, F.; Plitz, I.; Grugeon, S.; Laruelle, S.; Dollé, M.; Gozdz, A.S.; Tarascon, J.-M. Metal oxides as negative electrode materials in Li-ion cells. Electrochem. Solid-State Lett. **2002**, *5* (6), A115–A118.
18. Obrovac, M.N.; Dunlap, R.A.; Sanderson, R.J.; Dahn, J.R. The electrochemical displacement reaction of lithium with metal oxides. J. Electrochem. Soc. **2001**, *148* (6), A576–A588.
19. Delahaye-Vidal, A.; Beaudoin, B.; Figlarz, M. Textural and structural studies on nickel hydroxide electrodes. I. Crystallized nickel hydroxide materials submitted to chemical and electrochemical redox cycling. React. Solids **1986**, *2*, 223–233.
20. Poizot, P.; Laruelle, S.; Grugeon, S., Tarascon, J.-M. Rationalization of the low-potential reactivity of 3d-metal-based inorganic compounds toward Li. J. Electrochem. Soc. **2002**, *149* (9), A1212–A1217.
21. Pereira, N.; Balasubramanian, M.; Dupont, L.; McBreen, J.; Klein, L.C.; Amatucci, G.G. The electrochemistry of germanium nitride with lithium. J. Electrochem. Soc. **2003**, *150* (8), A1118–A1128.
22. Pereira, N.; Klein, L.C.; Amatucci, G.G. The electrochemistry of Zn_3N_2 and LiZnN. A lithium reaction mechanism for metal nitride electrodes. J. Electrochem. Soc. **2002**, *149* (3), A262–A271.
23. Shao-Horn, Y.; Osmialowski, S.; Horn, Q.C. Reinvestigation of lithium reaction mechanisms in FeS_2 pyrite at ambient temperature. J. Electrochem. Soc. **2002**, *149* (12), A1547–A1555.
24. Poizot, P.; Laruelle, S.; Grugeon, S.; Dupont, L.; Tarascon, J.-M. Searching for new anode materials for the Li-ion technology: Time to deviate from the usual path. J. Power Sources **2001**, *97–98*, 235–239.
25. Peled, E. The electrochemical behavior of alkali and alkaline earth metals in nonaqueous battery systems. The solid electrolyte interphase model. J. Electrochem. Soc. **1979,** 126, 2047–2051.
26. Rao, B.M.L.; Francis, R.W.; Christopher, H.A. Lithium–aluminum electrode. J. Electrochem. Soc. **1977,** 124, 1490–1492.
27. Nicholson, M.M. Lithium–magnesium electrodes in propylene carbonate. J. Electrochem. Soc. **1974,** 121, 734–738.
28. Huggins, R.A. Lithium alloy negative electrodes formed from convertible oxides. Solid State Ionics **1998**, *113–115*, 57–67.
29. Huggins, R.A. Lithium alloy negative electrodes. J. Power Sources **1999**, *81–82*, 13–19.
30. Anani, A.; Huggins, R.A. Multinary alloy electrodes for solid state batteries. J. Power Sources **1992**, *38*, 351–362.
31. Benedek, R.; Thackeray, M.M. Lithium reactions with intermetallic-compound electrodes. J. Power Sources **2002**, *110*, 406–411.
32. Thackeray, M.M.; Vaughey, J.T.; Johnson, C.S.; Kropf, A.J.; Benedek, R.; Fransson, L.M.L.; Edstrom, K. Structural considerations of intermetallic electrodes for lithium batteries. J. Power Sources **2003**, *113*, 124–130.
33. Wang, J.; Raistrick, I.D.; Huggins, R.A. Behavior of some binary alloys as negative electrodes in organic solvent-based electrolytes. J. Electrochem. Soc. **1986**, *133*, 457.
34. Yang, J.; Winter, M.; Besenhard, J.O. Small particle size multiphase Li-alloy anodes for lithium-ion batteries. Solid State Ionics **1996**, *90*, 281–287.
35. Yang, J.; Takeda, Y.; Imanishi, N.; Yamamoto, O. Ultrafine Sn and $SnSb_{0.14}$ powders for lithium storage matrices in lithium-ion batteries. J. Electrochem. Soc. **1999**, *146* (11), 4009–4013.
36. Courtney, I.A.; McKinnon, W.R.; Dahn, J.R. On the aggregation of tin in SnO composite glasses caused by the reversible reaction with lithium. J. Electrochem. Soc. **1999**, *146* (1), 59–68.
37. Idota, Y.; Kabuto, T.; Matsufuji, A.; Maekawa, Y.; Miyasaki, T. Tin-based amorphous oxides, a high-capacity lithium-ion storage material. Science **1997**, *276*, 1395–1397.
38. Courtney, I.A.; Dahn, J.R. Key factors controlling the reversibility of the reaction of lithium with SnO_2 and Sn_2BPO_6 glass. J. Electrochem. Soc. **1997**, *144* (9), 2943–2948.
39. Courtney, I.A.; Dahn, J.R. Electrochemical and in situ x-ray diffraction studies of the reaction of lithium with tin oxide composites. J. Electrochem. Soc. **1997**, *144* (6), 2045–2052.
40. Wilson, A.M.; Dahn, J.R. Lithium insertion in carbons containing nanodispersed silicon. J. Electrochem. Soc. **1995**, *142* (2), 323–326.
41. Yang, J.; Wang, B.F.; Wang, K.; Liu, Y.; Xie, J.Y.; Wen, Z.S. Si/C composites for high capacity lithium storage materials. Electrochem. Solid-State Lett. **2003**, *6* (8), A154–A156.
42. Wang, C.S.; Wu, G.T.; Zhang, X.B.; Qi, Z.F.; Li, W.Z. Lithium insertion in carbon–silicon composite materials produced by mechanical milling. J. Electrochem. Soc. **1998**, *145* (8), 2751–2758.
43. Mao, O.; Dunlap, R.A.; Dahn, J.R. Mechanically alloyed Sn–Fe(–C) powders as anode materials for Li-ion batteries. I. The Sn_2Fe–C system. J. Electrochem. Soc. **1999**, *146* (2), 405–413.

44. Mao, O.; Dahn, J.R. Mechanically alloyed Sn–Fe-(–C) powders as anode materials for Li-ion batteries. II. The Sn–Fe system. J. Electrochem. Soc. **1999**, *146* (2), 414–422.
45. Mao, O.; Dahn, J.R. Mechanically alloyed Sn–Fe-(–C) powders as anode materials for Li-ion batteries. III. Sn_2Fe:$SnFe_3C$ active/inactive composites. J. Electrochem. Soc. **1999**, *146* (2), 423–427.
46. Beaulieu, L.Y.; Larcher, D.; Dunlap, R.A.; Dahn, J.R. Reaction of Li with grain-boundary atoms in nanostructured compounds. J. Electrochem. Soc. **2000**, *147* (9), 3206–3212.
47. Beaulieu, L.Y.; Dahn, J.R. The reaction of lithium with Sn–Mn–C intermetallics prepared by mechanical alloying. J. Electrochem. Soc. **2000**, *147* (9), 3237–3241.
48. Scrosati, B.; Croce, F.; Persi, L. Impedance spectroscopy study of PEO-based nanocomposite polymer electrolytes. J. Electrochem. Soc. **2000**, *147* (5), 1718–1721.
49. Persi, L.; Croce, F.; Scrosati, B.; Plichta, E.; Hendrickson, M.A. Polyethylene oxide-based, nanocomposite electrolytes as improved separators for rechargeable lithium polymer batteries. The Li/$LiMn_3O_6$ case. J. Electrochem. Soc. **2002**, *149* (2), A212–A216.
50. Appetecchi, G.B.; Croce, F.; Persi, L.; Ronci, F.; Scrosati, B. Transport and interfacial properties of composite polymer electrolytes. Electrochim. Acta **2000**, *45*, 1481–1490.
51. Du Pasquier, A.; Laforgue, A.; Simon, P.; Amatucci, G.G.; Fauvarque, J.-F. A nonaqueous asymmetric hybrid $Li_4Ti_5O_{12}$/poly(fluorophenylthiophene) energy storage device. J. Electrochem. Soc. **2002**, *149* (3), A302–A306.

Palladium Nanoclusters: Preparation and Synthesis

Kiyotomi Kaneda
Kwang-Min Choi
Tomoo Mizugaki
Kohki Ebitani
Osaka University, Osaka, Japan

INTRODUCTION

The chemistry of polynuclear metal clusters in the nanometer scale has become a central issue in materials science because of their unique functions as electronic, optical, magnetic, and catalytic devices, which bring new technologies in many industrial areas.[1–5] At present, there is some progress being made in the preparation of metal nanoclusters.[6–15] They have been synthesized by the chemical reduction of metal salts in the presence of stabilizing ligands such as surfactants,[6,7] organic polymers,[8–10] and organic bases,[11–13] which preserve the metal clusters from the agglomeration of the clusters themselves. However, the above methods hardly produce metal nanoclusters with a high degree of size dispersity.

Transition metals have been extensively used as catalysts for many organic reactions, where the oxidation state of the metal plays an important role in attaining highly selective reactions.[16–19] A multiple interaction of substrate molecules with surface metal atoms of transition metal nanoclusters would offer the possibility of unprecedented catalytic reactions. Furthermore, a creation of specific surface ensemble sites consisting of zero-valent metals and metal cations is a promising strategy for designing highly functionalized nanocluster catalysts based on cooperative action among the multiple metal species.

The purpose of this article is to develop a new synthetic protocol of monodispersed palladium (Pd) nanoclusters at the nanometer-scale precision as high-performance catalysts. We describe here the novel synthesis of Pd nanoclusters with a high degree of size dispersity, where the standard deviation of the mean diameter (σ/d) is less than 10%, by treatment of the small Pd cluster with metal nitrates such as $Cu(NO_3)_2$ under an O_2 atmosphere. This method also yielded mixed-valence states with Pd^0 and cationic Pd species on the cluster surface.[20] An advantage of this synthetic method is that the particle size and the surface oxidation state of monodispersed Pd nanoclusters can be controlled by selecting the amount of the metal nitrates and by varying the preparation time of the clusters. Furthermore, the above Pd nanoclusters are applicable to highly efficient and selective heterogeneous catalysts for liquid-phase oxidations such as acetoxylation of toluene, alcohol oxidation, and the Wacker oxidation under an atmospheric pressure of O_2.[20–22] The unique catalyses of the Pd nanoclusters are attributed to the cooperative action between zero-valent Pd and Pd cations on the surface ensemble Pd sites.

PREPARATION OF NANOCLUSTERS

A representative example is the synthesis of the monodispersed Pd nanocluster with a mean diameter of 38 Å. A solution of $Pd(OAc)_2$ (0.20 g, 0.89 mmol) in AcOH (20 mL) was stirred at 50°C for 2 hr under continuous CO flow, yielding 0.12 g of $Pd_4(CO)_4(OAc)_4 \cdot 2AcOH$ (PCA) as a yellow precipitate.[23] The obtained PCA (0.12 g, 0.155 mmol, Pd 0.62 mmol) was stirred in AcOH (2.5 mL) in the presence of 1,10-phenanthroline (phen) (0.0614 g, 0.31 mmol) at room temperature for 30 min in air to give a dark brown solution of $Pd_4phen_2(CO)_2(OAc)_4$[24] [anal. calc. for $Pd_4C_{34}H_{28}N_4O_{10}$ (%): C, 37.9; H, 2.6; N, 5.2; found (%): C, 34.1; H, 2.7; N, 5.2].[a] Then, 0.0038 g of $Cu(NO_3)_2 \cdot 3H_2O$ (0.0155 mmol) was added to the above solution and heated at 90°C under an atmospheric O_2. After 15 min, a black solid was precipitated with the consumption of 0.067 mmol O_2 (O_2/Pd=0.108 mol/mol).[b] The precipitate was washed with AcOH for several times and dried in vacuo to yield the Pd nanoclusters (0.032 g, ca. 30% yield based on Pd) [anal. calc. for $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$ (%): C, 3.28; H, 0.41; N, 1.91; found (%): C, 3.21; H, 0.23; N, 1.34; X-ray photon spectroscopy (XPS): Pd $3d_{5/2}$=334.5–335.2 eV, O 1s=531.5 eV, N 1s=399.4 eV. The XPS peak positions are referred to C 1s at 284.6 eV (atomic ratio by

[a]An optimum mole ratio of Pd to phen in the synthesis of giant Pd clusters was found to be 2. The use of $Pd_4phen_4(CO)_2(OAc)_4$ instead of $Pd_4phen_2(CO)_2(OAc)_4$ did not afford giant Pd clusters.

[b]The amount of O_2 uptake during the preparation of the Pd nanoclusters was volumetrically measured using a gas buret by trapping evolved CO_2 in a cold trap at −120°C.

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013764

$$90[Pd_4phen_2(CO)_2(OAc)_4] + 9[Cu(NO_3)_2 \cdot 3H_2O] + 172\ AcOH + 137\mathbf{O_2}$$
$$\rightarrow 0.05[Pd_{2060}(NO_3)_{360}(OAc)_{360}\mathbf{O_{80}}] + 180[Pdphen(OAc)_2]$$
$$+ 77[Pd(OAc)_2] + 9[Cu(OAc)_2] + 113\mathbf{H_2O} + 180\mathbf{CO_2}$$

Scheme 1 Formation of the $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$ cluster.

XPS: C/Pd=0.6, N/Pd=0.12). Copper signals were not appreciable on the clusters.].[c] Scheme 1 represents the formation of the Pd nanoclusters. During the preparation of the Pd nanoclusters, molecular oxygen was consumed to form CO_2 and H_2O, and the oxygen species on the cluster surface.

A typical example for the immobilization of the Pd clusters on the metal oxides is described for the Pd nanocluster/TiO_2. TiO_2 (JRC-TIO-2, 1.30 g) was added into the acetic acid solution of the $Pd_4phen_2(CO)_2(OAc)_4$ and $Cu(NO_3)_2 \cdot 3H_2O$, and then stirred at 90°C under an atmospheric O_2 (3 wt.% as Pd cluster).[d] The solid compound of 1.4 g was obtained after filtration, washing with acetic acid and drying under a vacuum.

In the X-ray diffraction (XRD) pattern of the Pd nanoclusters, the three peaks are observed at 40°, 46°, and 68°, which correspond to the {111}, {200}, and {220} planes of a face-centered cubic (fcc) lattice, respectively. The mean crystalline size of the Pd nanoclusters, calculated from the line broadening of the peak at 40° using the Scherrer equation, is ca. 35 Å. The particle size distribution and the ordering of Pd atoms of the Pd clusters were estimated by field emission scanning electron microscopy (FE-SEM) and high-resolution transmission electron microscopy (HR-TEM) of the Pd clusters immobilized on the TiO_2 surface because the Pd nanoclusters could be immobilized on a surface of TiO_2 without a change in their original cluster size and the local ordering of the Pd atoms on the cluster surface. As shown in Fig. 1a, the FE-SEM image of the Pd clusters on the TiO_2 surface reveals that the Pd nanoclusters are well dispersed with a quite narrow size distribution throughout the TiO_2 surface. The mean diameter (d) of the Pd clusters is 38 Å with a standard deviation (σ) of 2.1 Å, where the σ/d value is 6%. This particle size corresponds to that of the giant eight-shell metal cluster composed of 2060 Pd atoms as the magic number.[25] To the best of our knowledge, the Pd nanoclusters prepared by our method using metal nitrates have the narrowest size distribution of Pd nanoparticles obtained by chemical and electrochem ical methods.[6–15]

Fig. 2 represents the HR-TEM image of the Pd nanocluster immobilized on the TiO_2 surface. A regular arrangement of Pd atoms on the Pd nanocluster surface can be clearly observed, with 17 Pd atoms along the {111} planes. The Pd K-edge X-ray absorption fine structure (XAFS) measurement of the Pd nanoclusters indicated that the inner part is composed of a metallic Pd atom. However, the surface analysis of the above Pd nanoclusters by XPS proved the presence of both cationic and metallic Pd species. Furthermore, the treatment of the Pd nanoclusters with a CO molecule at 0°C led to both CO_2 evolution and CO adsorption on metallic Pd; the cationic Pd species are reduced to Pd^0 during exposure to CO. The evolution of CO_2 ascertains the presence of oxygen species on the surface of the Pd nanoclusters. Because the reactivity of the surface oxygen species was high enough to oxidize a CO molecule even at 0°C, the Pd sites responsible for the above CO oxidation would be Pd_2O species on the cluster surface.[e] The ratio of the Pd_2O species to the surface Pd atoms of the monodispersed Pd nanoclusters with a diameter of 38 Å, evaluated from the amounts of the formed CO_2 and the CO uptake, is 23%.

A possible arrangement of the Pd atoms of the eight-shell Pd cluster with an fcc cuboctahedral shape is illustrated in Scheme 2a, where the Pd^0 and Pd^+ cation species are situated on the face of the nanocluster. The fraction of Pd^+ (Pd_2O) species of the surface Pd atoms was 23%, whereas those of Pd^{2+} and Pd^0 species were 56% and 20%, respectively.[20] This accords with the composition of the Pd nanoclusters, $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$; the surface-divalent Pd cations were preserved at corner/edge sites by anions of NO_3^- and OAc^-, where oxygen species were bound to the two monovalent Pd ions on the face of the nanocluster.

The use of $Fe(NO_3)_3$ instead of $Cu(NO_3)_2$ could also afford the Pd nanoclusters having a narrow size distribution with a mean diameter and standard deviation ($d \pm \sigma$) of 38±4.4 Å (σ/d=11%), where the fraction of the surface Pd_2O species exhibited the maximum of 22%. The composition of the eight-shell Pd nanoclusters was $Pd_{2060}(NO_3)_{250}(OAc)_{470}O_{80}$. The ratio of OAc^- to NO_3^- anions was different from that obtained for the eight-shell Pd clusters using $Cu(NO_3)_2$. However, the total number of anionic species situated on the cluster surface was constant, irrespective of the metal nitrates.

Furthermore, we found that the selection of the $Cu(NO_3)_2$/PCA molar ratio can control the particle size of the Pd nanoclusters.[20] When the $Cu(NO_3)_2$/PCA

[c]The Pd(phen)$(OAc)_2$ complex was formed in the residual solution.

[d]TiO_2 (JRC-TIO-2) was supplied from the Catalysis Society of Japan (anatase, BET surface area: 14 $m^2\ g^{-1}$), and is characterized by its low acidity and basicity compared with other TiO_2 reference samples.

[e]A Pd^+ species has been already reported on palladium-loaded silico-aluminophosphate molecular sieves. (See Ref. [26].)

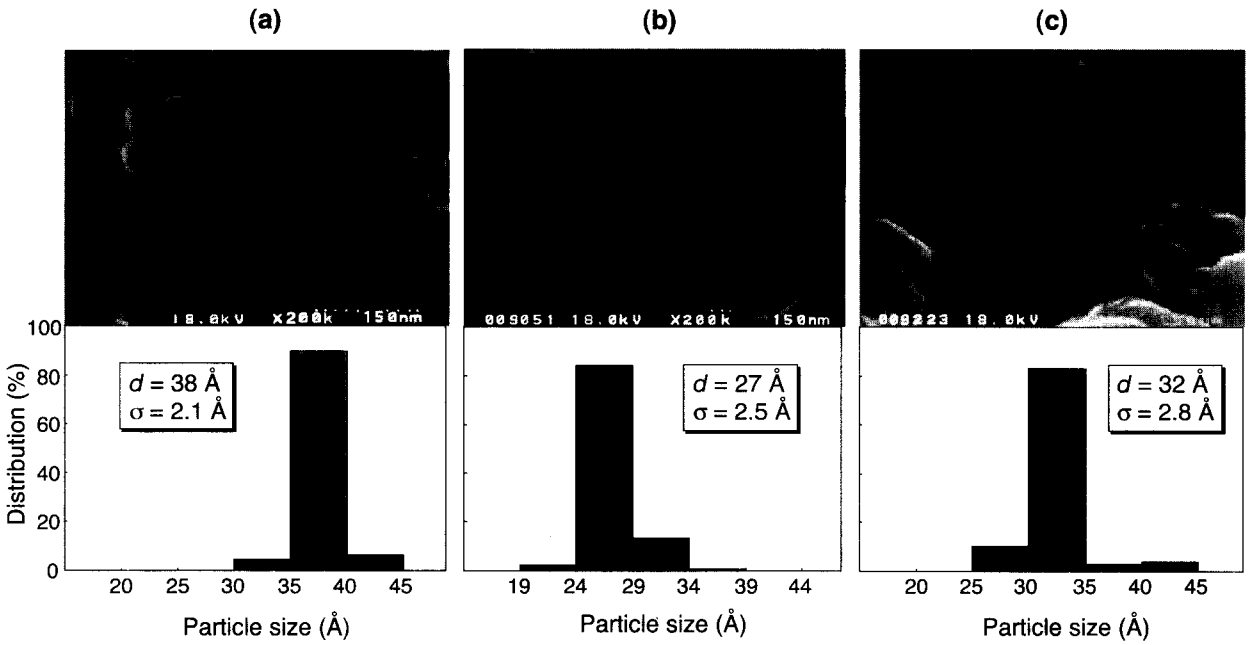
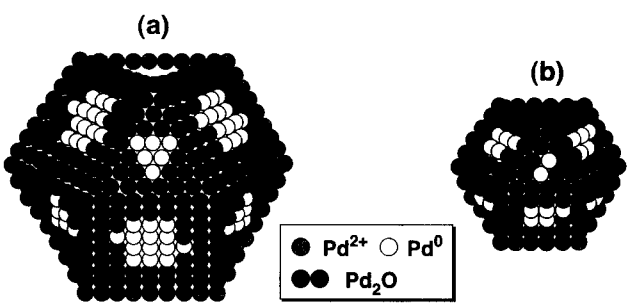


Fig. 1 The FE-SEM photograph and the particle size distribution of the Pd nanoclusters prepared with the $Cu(NO_3)_2$/PCA molar ratio of (a) 0.10 (15 min), (b) 0.05 (8 min), and (c) 0.10 (10 min). The values in parentheses are the preparation times of the cluster.

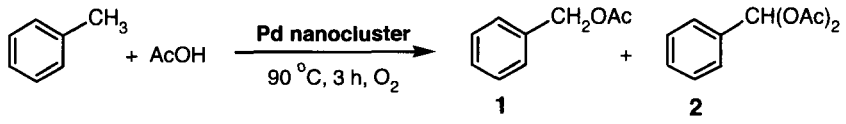
molar ratio was 0.05, the five-shell Pd nanoclusters with the σ/d value of 9% (σ=2.5 Å, d=27 Å) were obtained. A typical FE-SEM image and the particle size distribution are shown in Fig. 1b. The fraction of the Pd^+ species of the surface Pd atoms became 7% when the $Cu(NO_3)_2$/PCA molar ratio was 0.05. Scheme 2b shows a proposed structure of the five-shell Pd nanocluster composed of 560 Pd atoms whose composition is $Pd_{560}(NO_3)_{100}(OAc)_{250}O_{10}$. It is notable that the treatment of $Pd_4phen_2(CO)_2(OAc)_4$ with $Cu(NO_3)_2 \cdot 3H_2O$ in a shorter preparation time (e.g., ~10 min) afforded Pd nanoclusters with a σ/d value of 9% (σ=2.8 Å, d=32 Å), as depicted in Fig. 1c. This particle size might correspond with six-shell or seven-shell clusters.[12]

The above shell-selective synthesis of monodispersed Pd clusters allows us to consider the feature of nanocluster formation as follows. The metal nitrates promote the disproportionation of monovalent Pd ions of $Pd_4phen_2(CO)_2(OAc)_4$ complex to yield Pd^0 and Pd^{2+} cation species. The Pd^0 species aggregate to form relatively small Pd assemblies. Because the amounts of Pd^0 and Pd^{2+} species increase with increasing amount of the metal nitrates and duration of the treatment, the small assembly of Pd^0 species is overspread with generated Pd^0 species,

Fig. 2 The HR-TEM image of the eight-shell Pd nanocluster immobilized on the TiO_2 surface.



Scheme 2 Possible structure models of (a) the eight-shell Pd nanocluster, $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$, and (b) the five-shell Pd nanocluster, $Pd_{560}(NO_3)_{100}(OAc)_{250}O_{10}$, with an fcc cuboctahedral shape.



Scheme 3 Oxidative acetoxylation of toluene in the presence of molecular oxygen catalyzed by the eight-shell Pd nanoclusters.

which lead to the growth of a larger Pd^0 nanocluster. Subsequently, the Pd^0 atoms located at corner/edge sites of the larger cluster surface are oxidized by the metal nitrate to give Pd^{2+} cations, which are stabilized by anionic ligands such as OAc^- and NO_3^-. On the other hand, the Pd_2O surface species might be formed by the oxidation of the Pd^0 ions by O_2. An additional increase in the amount of the metal nitrates resulted in a deep oxidation of the surface Pd species and also an exfoliation of Pd^{2+} ions from the cluster surface, which resulted in the formation of smaller metallic clusters together with decomposition into the monomeric Pd^{2+} complex. Indeed, both the particle size and the surface oxidation state of the Pd cluster exhibited maxima at a $Cu(NO_3)_2$/PCA molar ratio of 0.10.[20]

The quite narrow size distribution of the present Pd nanoclusters could be ascribed to the strong ability of the metal nitrate for the disproportionation of monovalent Pd ions of the $Pd_4phen_2(CO)_2(OAc)_4$ complex that leads to a formation of a relatively small Pd assembly through a fast nucleation step of the Pd^0 species. Furthermore, the partial oxidation of surface Pd^0 species at the larger cluster surface with metal nitrate retards the successive growth of the nanoclusters.

Conclusively, the treatment of palladium carbonyl acetate (PCA) cluster with $Cu(NO_3)_2$ in the presence of phen under an O_2 atmosphere afforded monodispersed Pd nanoclusters containing both Pd^0 and cationic Pd species on the surface. The novel preparation method described here enables a "shell-selective" synthesis of Pd nanoclusters, keeping the high monodispersity as well as a tuning of the surface fraction of Pd^0, Pd^+, and Pd^{2+} species. Unique catalyses of the Pd nanoclusters in the acetoxylation of toluene, alcohol oxidation, and the Wacker oxidation are described below.

APPLICATION

Palladium is among the most important catalysts in organic synthesis.[17,18,27,28] Our monodispersed Pd nanoclusters will offer the possibility of unprecedented catalytic reactions based on specific surface ensemble Pd sites within a regular arrangement of multiple Pd species. We applied the monodispersed Pd nanoclusters to heterogeneous catalysts for oxidative functional transformations in liquid phase under an atmospheric O_2 pressure.

Acetoxylation of Toluene

Acetoxylation is one of the powerful methods to functionalize various kinds of C—H bonds in hydrocarbons; the acetoxyl group can be facilely transformed into other functional ones. Generally, the acetoxylation of toluene occurs by using a homogeneous $Pd(OAc)_2$ complex in an acetic acid under an O_2 atmosphere to give benzyl acetates, where additives of KOAc and $Sn(OAc)_2$ are indispensable.[29] In heterogeneous Pd catalysts, Pd particles on charcoal or SiO_2, as a support interacting with excess amounts of Sn(2-ethyl-hexanoate)$_2$ and KOAc, exhibit catalytic activity.[30,31] It is said that the efficient acetoxylation by Pd catalysts cannot be attained without any additives.

Our monodispersed Pd nanoclusters acted as heterogeneous catalysts for the oxidative acetoxylation of toluene to benzyl acetates using molecular oxygen (Scheme 3).[20] Among the Pd nanoclusters prepared using $Cu(NO_3)_2$, the maximum catalytic activity was obtained in the case of the eight-shell Pd nanoclusters having the highest fraction of Pd_2O and the divalent Pd species. It is interesting to note that the monodispersed Pd nanoclusters can efficiently catalyze the oxidative acetoxylation of toluene without any additives such as Sn(II) compounds, KOAc, and reoxidizing reagents. Furthermore, the eight-shell Pd nanoclusters showed much higher catalytic activity than conventional Pd catalysts [e.g. Pd/carbon, Pd/Al_2O_3, and $Pd(OAc)_2$]. The eight-shell Pd nanoclusters immobilized on the TiO_2 surface also had high catalytic activity for the acetoxylation of toluene.[f] These results are summarized in Table 1.

The high catalytic activity of the Pd nanoclusters is ascribed to an ensemble of the cationic Pd species and Pd^0 on the cluster surface. Vide supra, in the conventional oxidative acetoxylation catalyzed by Pd catalyst systems, an addition of potassium and tin compounds is indispensable. The Sn^{2+} additives promote the reduction step of Pd^{2+} to Pd^0, and the resulting Sn^{4+} oxide may serve as an oxygen reservoir in oxygen-poor regimes. The potassium

[f]The particle size of the Pd clusters did not change appreciably after the acetoxylation reaction. During the acetoxylation. no Pd leaching was observed by the ICP method, whose detection limit is 24 ppb.

Table 1 Acetoxylation of toluene with various Pd catalysts in the presence of molecular oxygen[a]

Catalyst	Conversion (%)	Total yield of 1+2 (%)[b]	1:2
Pd nanocluster[c]	96	94	4.1:1
Pd nanocluster/TiO_2	93	88	5.0:1
5% Pd/carbon	19	Trace	–
5% Pd/Al_2O_3	21	0	–
$Pd(OAc)_2$	0	0	–

[a]Reaction conditions: Pd 0.25 mmol, toluene 5 mmol, AcOH 4 mL, 90°C, 3 hr, O_2 atmosphere.
[b]Yields of products were determined by gas chromatography (GC) analysis using an internal standard technique.
[c]Pd (0.29 mmol).

renders Pd particles more electropositive, which becomes susceptible to interaction with an electron-donating substrate (e.g., toluene). Because the cationic Pd atoms already exist on the surface of the Pd cluster in the present catalyst system, the addition of potassium compounds is not required to facilitate the interaction with toluene. Furthermore, tin compound is also unnecessary because the Pd_2O species acted as an oxygen reservoir through an easy interconversion between Pd^+ and Pd^0 species.

Alcohol Oxidation

Recently, much attention has focused on Pd-catalyzed oxidation of alcohols into the corresponding carbonyl compounds using molecular oxygen as an environment-friendly oxidant.[32–37] In this section, we explore the catalytic potential of the monodispersed Pd nanoclusters in the aerobic oxidation of alcohols.[21]

Results of the aerobic oxidation of various alcohols catalyzed by the $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$ nanoclusters are displayed in Table 2 together with those using the $Pd_{561}phen_{60}(OAc)_{180}$ cluster[38] and $Pd(OAc)_2$-pyridine[35] catalytic systems. Our Pd nanoclusters efficiently catalyzed the oxidation of *primary* aromatic allylic alcohols to the corresponding aldehydes using molecular oxygen (entries 1, 4, and 5). It should be noted that the present Pd nanoclusters showed higher reactivity for *primary* aromatic allylic alcohols than for aliphatic ones (e.g., 3-methyl-2-buten-1-ol and 2-hexen-1-ol, entries 6 and 8). Furthermore, the oxidation of *secondary* aromatic allylic alcohols proceeded slowly (entry 10). Interestingly, benzylic alcohols were oxidized more slowly than allylic alcohols (entries 14 and 17). This high reactivity for allylic alcohols was also exemplified in a competitive reaction of an equimolar mixture of cinnamyl alcohol and benzyl alcohol; cinnamaldehyde was selectively obtained in 92% yield, together with only 1% of benzaldehyde for 3 hr (Fig. 3).

Generally, divalent Pd catalysts are prone facilely to oxidize benzyl alcohols, but are not suitable for the oxidation of allylic alcohols because of an irreversible coordination of allylic alcohols to the Pd species (entry 3 vs. entry 16).[35,39] The difference between the present Pd nanoclusters and conventional Pd^{2+} complexes lies in the high reactivity for allylic alcohols over benzylic alcohols. As compared with the $Pd_{561}phen_{60}(OAc)_{180}$ cluster, having Pd^+ cation species throughout the surface,[38] the present Pd nanoclusters preferred aromatic allylic alcohols to aliphatic ones. Cinnamyl alcohol was completely oxidized only for 1 hr, whereas the oxidation of 3-methyl-2-buten-1-ol required 5 hr (entry 1 vs. entry 6). In contrast, using the $Pd_{561}phen_{60}(OAc)_{180}$ catalyst, the oxidation of both allylic alcohols proceeded in a similar rate (entry 2 vs. entry 7).

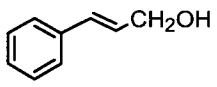
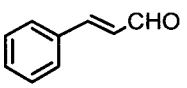
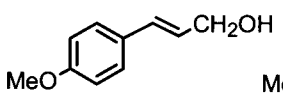
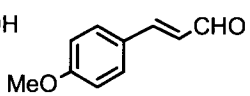
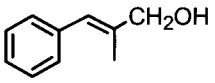
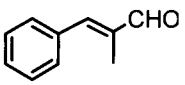
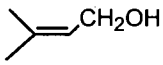
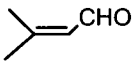
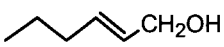
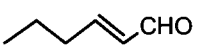
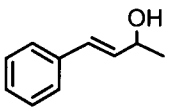
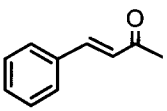
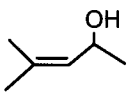
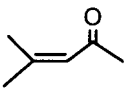
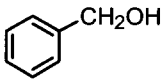
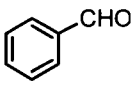
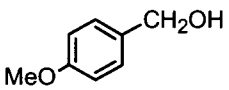
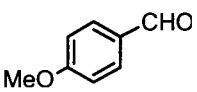
The highest catalytic activity was obtained for the $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$ nanoclusters, which have the highest percentage of the Pd^+ species among the surface Pd atoms.[21] Therefore, in a similar fashion of the above acetoxylation, we considered that the highly selective oxidation of aromatic alcohols by the Pd nanoclusters could be explained by the unique surface ensemble Pd sites; a π-bond interaction between aromatic allylic alcohols and the cationic Pd^{2+} species. The high selectivity for the oxidation of allylic alcohols when compared with benzyl alcohol might be related to a coordination of the C=C bond with a Pd cation (e.g., Pd^{2+}). Correspondingly, the distance between the Pd cation and Pd^0 fits with that found for the distance between a terminal β-carbon of a C=C bond and an oxygen atom of the OH of allylic alcohols.[g]

Wacker Oxidation

The Wacker oxidation is a powerful method to synthesize methyl ketones from terminal olefins, catalyzed by an

[g]The EXAFS measurement of the Pd nanoclusters showed that the average distance between the Pd species was 2.75 Å. The distance between a terminal β-carbon of the C=C bond and oxygen atom of the OH function of cinnamyl alcohol was ca. 3 Å, calculated by the PM3 semiempirical method as implemented in MOPAC.

Table 2 Aerobic oxidation of various alcohols catalyzed by $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$ nanoclusters[a]

Entry	Substrate	Product	Catalyst	Time (hr)	Conversion (%)	Yield (%)[b]
1			$Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$	2	93	91
2			$Pd_{561}phen_{60}(OAc)_{180}$	1	100	94
3			$Pd(OAc)_2$-pyridine[c]	2	46	35
4			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	2	100	91
5			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	4	91	87
6			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	5	90	83
7			$Pd_{561}phen_{60}(OAc)_{180}$	1	100	89
8			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	5	66	49[d]
9			$Pd_{561}phen_{60}(OAc)_{180}$	1	100	79
10			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	14	94	80
11			$Pd_{561}phen_{60}(OAc)_{180}$	4	47	36
12			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	24	<1	Trace
13			$Pd_{561}phen_{60}(OAc)_{180}$	24	8	1
14			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	14	95	92
				4	45	44
15			$Pd_{561}phen_{60}(OAc)_{180}$	24	78	24
16			$Pd(OAc)_2$-pyridine[c]	2	100	100
17			$Pd_{2060}(NO_3)_{360}(OAC)_{360}O_{80}$	14	95	91
				4	53	50
18			$Pd(OAc)_2$-pyridine[c]	2	97	95

[a]Reaction conditions: Pd 0.05 mmol, substrate 2 mmol, AcOH 4 mL, 60°C, O_2 atmosphere.
[b]Yields of aldehydes and ketones were determined by GC analysis using an internal standard technique.
[c]Substrate 2 mmol, $Pd(OAc)_2$ 0.05 mmol, pyridine 0.2 mmol, MS3A 0.5 g, toluene 4 mL, 80°C, O_2 atmosphere.
[d]Fifteen percent of 2-hexenyl acetate was formed.

aqueous solution of Pd^{2+} salt combined with Cu^{2+} and HCl under aerobic conditions.[27] The HCl is required to achieve a favorable reoxidation of Pd^0 by Cu^{2+} ions, and to prevent the aggregation of transient atomic Pd^0 species to inactive Pd precipitates. An acidic chloride medium not only corrodes the reactor wall, but also leads to the formation of chlorinated by-products. Hence much effort has been devoted to the development of HCl-free Wacker systems.[40–42] However, there are few heterogeneous catalysts for the Wacker oxidation of unreactive higher terminal olefins in liquid phase.[43–46] We found that the Pd nanoclusters immobilized on a TiO_2 surface acted as a highly active and recyclable catalyst for the liquid-phase Wacker oxidation of higher terminal olefins in

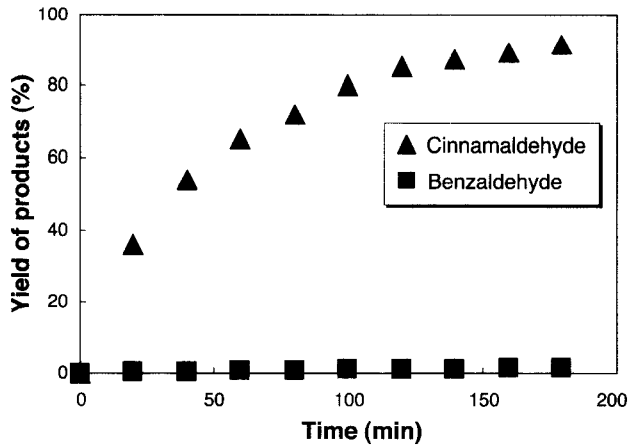
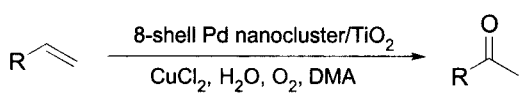


Fig. 3 Competitive oxidation of an equimolar mixture of cinnamyl alcohol and benzyl alcohol. Reaction conditions: $Pd_{2060}(NO_3)_{360}(OAc)_{360}O_{80}$ 0.0063 g (Pd; 0.05 mmol), cinnamyl alcohol 2 mmol, benzyl alcohol 2 mmol, AcOH 5 mL, 60 °C, O_2 atmosphere.

N,N-dimethylacetamide (DMA) solvent without acid additives (Scheme 4).[22]

The eight-shell Pd nanoclusters showed high catalytic activity for the oxidation of 1-decene in the presence of water and $CuCl_2 \cdot 2H_2O$ under O_2 at atmospheric pressure to afford 2-decanone selectively. The catalytic activity of the eight-shell Pd nanoclusters was higher than that of the five-shell $Pd_{560}(NO_3)_{100}(OAc)_{250}O_{10}$ clusters. In the present oxidation, the Pd nanoclusters immobilized on the TiO_2 could maintain the high catalytic activity and selectivity that make the workup procedure strikingly simple. Water, O_2, and $CuCl_2 \cdot 2H_2O$ were necessary to obtain high yields of 2-decanone. Only three equivalents of $CuCl_2$ to Pd were sufficient for achieving high catalytic activity. It is notable that the Pd nanocluster catalysts enable the Wacker oxidation in liquid phase under acid-free conditions.

Terminal olefins such as 1-hexene, 1-octene, 1-decene, 1-dodecene, vinylcyclohexane, and *n*-butyl vinyl ether were selectively oxidized by the immobilized Pd nanocluster catalysts to give the corresponding methyl ketones and *n*-butyl acetate, respectively, in high yields within 3 hr (Table 3). Furthermore, the spent catalyst was easily separated from the reaction mixture by filtration and could be reused with retention of high activity and



Scheme 4 The Wacker oxidation of terminal olefins catalyzed by eight-shell Pd nanoclusters immobilized on the TiO_2 surface.

Table 3 The Wacker oxidation of various terminal olefins catalyzed by eight-shell Pd nanoclusters/TiO_2[a]

Substrate	Yield (%)[b]	Selectivity (%)
1-Hexane	94[c]	98
1-Octene	84	98
1-Decene	88	97
1-Dodecene	83	98
n-Butyl vinyl ether	91	100
Vinylcyclohexane	92[d]	100

[a]Reaction conditions: Pd 0.01 mmol, substrate 1 mmol, $CuCl_2 \cdot 2H_2O$ 0.03 mmol, *N,N*-dimethylacetamide 4 mL, H_2O 0.5 mL, 80°C, 2 hr, O_2 atmosphere.
[b]Yields of products were determined by GC analysis using an internal standard technique.
[c]50°C.
[d]60°C, 3 hr.

selectivity. However, no oxidation occurred when the filtrate was further reacted. The present Wacker oxidation proceeds via the Pd^{2+} species on the Pd nanocluster surface. The use of even small amounts of $CuCl_2$ efficiently promoted the reoxidation of Pd^0 to Pd^{2+} species, which might be ascribed to a cooperative action of Pd^{2+}, Pd^+, and Pd^0 species on the nanocluster surface.

CONCLUSION

This study demonstrated the "shell-selective" synthesis and characterization of monodispersed Pd nanoclusters containing Pd^0, Pd^+, and Pd^{2+} species on the surface. Each surface Pd species (e.g., Pd^0, Pd^+, and Pd^{2+}) does not act individually as a catalytic site, but works cooperatively as a "trio" on the same cluster surface to attain unique catalytic performances for functional transformations under an O_2 atmosphere. Furthermore, the Pd nanoclusters could be immobilized on a neutral TiO_2 surface with retention of their original size and unique catalysis. The monodispersed Pd nanoclusters provides a new strategy to design highly *selective* Pd catalysts in many organic syntheses aiming at "green and sustainable chemistry".[47]

REFERENCES

1. Toshima, N.; Shiraishi, Y.; Teranishi, T.; Miyake, M.; Tominaga, T.; Watanabe, H.; Brijoux, W.; Bönnemann, H.; Schmid, G. Various ligand-stabilized metal nanoclusters as homogeneous and heterogeneous catalysts in the liquid phase. Appl. Organomet. Chem. **2001**, *15*, 178–196.

2. Rao, C.N.R.; Kulkarni, G.U.; Thomas, P.J.; Edwards, P.P. Metal nanoparticles and their assemblies. Chem. Soc. Rev. **2000**, *29*, 27–35.
3. Aiken, J.D., III; Finke, R.G. A review of modern transition-metal nanoclusters: Their synthesis, characterization, and applications in catalysis. J. Mol. Catal., A Chem. **1999**, *145*, 1–44.
4. Moiseev, I.I.; Vargaftik, M.N. Pd cluster catalysis: A review of reactions under anaerobic conditions. New J. Chem. **1998**, 1217–1227.
5. Schmid, G. Large clusters and colloids. Metals in the embryonic state. Chem. Rev. **1992**, *92*, 1709–1727.
6. Bradley, J.S.; Tesche, B.; Busser, W.; Masse, M.; Reetz, M.T. Surface spectroscopic study of the stabilization mechanism for shape-selectively synthesized nanostructured transition metal colloids. J. Am. Chem. Soc. **2000**, *122*, 4631–4636.
7. Bönnemann, H.; Braun, G.; Brijoux, W.; Brinkmann, R.; Tilling, A.S.; Seevogel, K.; Siepen, K. Nanoscale colloidal metals and alloys stabilized by solvents and surfactants. Preparation and use as catalyst precursors. J. Organomet. Chem. **1996**, *520*, 143–162.
8. Thomas, P.J.; Kulkarni, G.U.; Rao, C.N.R. Magic nuclearity giant clusters of metal nanocrystals formed by mesoscale self-assembly. J. Phys. Chem., B **2001**, *105*, 2417–2515.
9. Teranishi, T.; Miyake, M. Size control of palladium nanoparticles and their crystal structures. Chem. Mater. **1998**, *10*, 594–600.
10. Toshima, N. Nanostructured metal clusters in polymeric field as a model of artificial enzyme. Supramol. Sci. **1998**, *5*, 395–398.
11. Vargaftik, M.N.; Zagorodnikov, V.P.Stolarov, I.P.; Moiseev, I.I.; Likholobov, V.A.; Kotchubey, D.I.; Chuvilin, A.L.; Zaikovsky, V.I.; Zamaraev, K.I.; Timofeeva, G.I. A novel giant palladium cluster. J. Chem. Soc., Chem. Commun. **1985**, 937–939.
12. Schmid, G.; Harms, M.; Malm, J.-O.; Bovin, J.-O.; van Ruitenbeck, J.; Zandbergen, H.W.; Fu, W.T. Ligand-stabilized giant palladium clusters: Promising candidates in heterogeneous catalysis. J. Am. Chem. Soc. **1993**, *115*, 2046–2048.
13. Lin, Y.; Finke, R.G. Novel polyoxoanion- and Bu_4N^+-stabilized, isolatable, and redissolvable, 20–30 Å $Ir_{300-900}$ nanoclusters: The kinetically controlled synthesis, characterization, and mechanism of formation of organic solvent-soluble, reproducible size, and reproducible catalytic activity metal nanoclusters. J. Am. Chem. Soc. **1994**, *116*, 8335–8353.
14. Reetz, M.T.; Helbig, W. Size-selective synthesis of nanostructured transition metal clusters. J. Am. Chem. Soc. **1994**, *116*, 7401–7402.
15. Reetz, M.T.; Westermann, E. Phosphine-free palladium-catalyzed coupling reactions: The decisive role of Pd nanoparticles. Angew. Chem., Int. Ed. Engl. **2000**, *39*, 165–168.
16. Tsuji, J. *Organic Synthesis with Palladium Compounds*; Springer-Verlag: Berlin, 1980.
17. Heck, R.F. *Palladium Reagents in Organic Syntheses*; Academic Press: London, 1985.
18. Jia, C.; Kitamura, T.; Fujiwara, Y. Catalytic functionalization of arenes and alkanes via C—H bond activation. Acc. Chem. Res. **2001**, *34*, 633–639.
19. Sheldon, R.A.; Kochi, J.K. *Metal Catalyzed Oxidation of Organic Compounds*; Academic Press: New York, 1981.
20. Ebitani, K.; Choi, K.-M.; Mizugaki, T.; Kaneda, K. Novel preparation of palladium nanoclusters using metal nitrates and their catalysis for oxidative acetoxylation of toluene in the presence of molecular oxygen. Langmuir **2002**, *18*, 1849–1855.
21. Choi, K.-M.; Akita, T.; Mizugaki, T.; Ebitani, K.; Kaneda, K. Highly selective oxidation of allylic alcohols catalysed by monodispersed 8-shell Pd nanoclusters in the presence of molecular oxygen. New J. Chem. **2003**, *27*, 324–328.
22. Choi, K.-M.; Mizugaki, T.; Ebitani, K.; Kaneda, K. Nanoscale palladium cluster immobilized on a TiO_2 surface as an efficient catalyst for liquid-phase Wacker oxidation of higher terminal olefins. Chem. Lett. **2003**, *32*, 180–181.
23. Moiseev, I.I.; Stromnova, T.; Vargaftik, M.N.; Mazo, G.J.; Kuz'mina, L.G.; Struchkov, Y.T. New palladium carbonyl clusters: X-ray crystal structure of $[Pd_4(CO)_4(OAc)_4]\cdot(AcOH)_2$. J. Chem. Soc., Chem. Commun. **1978**, 27–28.
24. Moiseev, I.I. Pd_4 clusters: The sensitivity of the cluster metal core geometry to the surrounding ligands. J. Organomet. Chem. **1995**, *488*, 183–190.
25. Chini, P. Large metal carbonyl clusters. J. Organomet. Chem. **1980**, *200*, 37–61.
26. Choo, H.; Prakash, A.M.; Zhu, Z.; Kevan, L. Formation and adsorbate interactions of paramagnetic Pd(I) species in Pd(II)-exchanged NaK- and H-clinoptilolite. J. Phys. Chem., B **2000**, *104*, 3608–3615.
27. Tsuji, J. *Palladium Reagents and Catalysts Innovations in Organic Synthesis*; John Wiley and Sons: New York, 1998.
28. Trost, B.M. Atom economy—A challenge for organic synthesis: Homogeneous catalysis leads the way. Angew. Chem., Int. Ed. Engl. **1995**, *34*, 259–281.
29. Bryant, D.R.; McKeon, J.E.; Ream, B.C. Palladium-catalyzed synthesis of benzyl esters from methylbenzenes. J. Org. Chem. **1968**, *33*, 4123–4127.

30. Benazzi, E.; Mimoun, H.; Cameron, C.J. Heterogeneous catalyzed benzylic acetoxylation of methylated aromatic hydrocarbons. J. Catal. **1993**, *140*, 311–327.
31. Tanielyan, S.K.; Augustine, R.L. Pd/Sn catalyst for toluene acetoxylation. J. Mol. Catal. **1994**, *90*, 267–289.
32. Mori, K.; Yamaguchi, K.; Hara, T.; Mizugaki, T.; Ebitani, K.; Kaneda, K. Controlled synthesis of hydroxyapatite-supported palladium complexes as highly efficient heterogeneous catalysts. J. Am. Chem. Soc. **2002**, *124*, 11572–11573.
33. Steinhoff, B.A.; Fix, S.R.; Stahl, S.S. Mechanistic study of alcohol oxidation by the $Pd(OAc)_2/O_2$/DMSO catalyst system and implications for the development of improved aerobic oxidation catalysts. J. Am. Chem. Soc. **2002**, *124*, 766–767.
34. ten Brink, G.-J.; Arends, I.W.C.E.; Sheldon, R.A. Green, catalytic oxidation of alcohols in water. Science **2000**, *287*, 1636–1639.
35. Nishimura, T.; Onoue, T.; Ohe, K.; Uemura, S. Palladium(II)-catalyzed oxidation of alcohols to aldehydes and ketones by molecular oxygen. J. Org. Chem. **1999**, *64*, 6750–6755.
36. Peterson, K.P.; Larock, R.C. Palladium-catalyzed oxidation of primary and secondary allylic and benzylic alcohols. J. Org. Chem. **1998**, *63*, 3185–3189.
37. Kaneda, K.; Fujii, M.; Morioka, K. Highly selective oxidation of allylic alcohols to α,β-unsaturated aldehydes using Pd cluster catalysts in the presence of molecular oxygen. J. Org. Chem. **1996**, *61*, 4502–4503.
38. Ebitani, K.; Fujie, Y.; Kaneda, K. Immobilization of a ligand-preserved giant palladium cluster on a metal oxide surface and its novel heterogeneous catalysis for oxidation of allylic alcohols in the presence of molecular oxygen. Langmuir **1999**, *15*, 3557–3562.
39. Blackburn, T.F.; Schwartz, J. Homogeneous catalytic oxidation of secondary alcohols to ketones by molecular oxygen under mild conditions. J. Chem. Soc., Chem. Commun. **1977**, 157–158.
40. ten Brink, G.-J.; Arends, I.W.C.E.; Papadogianakis, G.; Sheldon, R.A. Catalytic conversions in water: Part 10. Aerobic oxidation of terminal olefins to methyl ketones catalysed by water soluble palladium complexes. Chem. Commun. **1998**, 2359–2360.
41. Smith, A.B., III; Cho, Y.S.; Friestad, G.K. Convenient Wacker oxidation with substoichiometric cupric acetate. Tetrahedron Lett. **1998**, *39*, 8765–8768.
42. Hosokawa, T.; Takano, M.; Murahashi, S.-I. The first isolation and characterization of a palladium–copper heterometallic complex bearing μ_4-oxo atom derived from molecular oxygen. J. Am. Chem. Soc. **1996**, *118*, 3990–3991.
43. De Vos, D.E.; Sels, B.F.; Jacobs, P.A. Immobilization of homogeneous oxidation catalysts. Adv. Catal. **2001**, *46*, 1–87.
44. Yokota, T.; Sakakura, A.; Tani, M.; Sakaguchi, S.; Ishii, Y. Selective Wacker-type oxidation of terminal alkenes and dienes using the Pd(II)/molybdovanadophosphate (NPMoV)/O_2 system. Tetrahedron Lett. **2002**, *43*, 8887–8891.
45. Kishi, A.; Higashino, T.; Sakaguchi, S.; Ishii, Y. Wacker-type oxidation of cyclopentene under dioxygen atmosphere catalyzed by $Pd(OAc)_2$/NPMoV on activated carbon. Tetrahedron Lett. **2000**, *41*, 99–102.
46. Tang, H.G.; Sherrington, D.C. Polymer-supported Pd(II) Wacker-type catalysts: II. Application in the oxidation of dec-1-ene. J. Catal. **1993**, *142*, 540–551.
47. *Handbook of Green Chemistry and Technology*; Clark, J., Macquarrie, D., Eds.; Blackwell Publishing: Oxford, 2002.

P

Phase Behavior of Nanoparticle Suspensions

P

S. Ramakrishnan
C. F. Zukoski
University of Illinois at Urbana-Champaign, Urbana, Illinois, U.S.A.

INTRODUCTION

The assembly of nanoparticles into interesting and useful structures requires the manipulation of particle interactions at the level of a few times the average thermal energy in the system such that the particles can sample an ensemble of states. When allowed to reach equilibrium, colloidal suspensions display phase behavior similar to that seen in molecular systems. The phases observed can be manipulated through alterations in particle density and particle surface chemistry. The richness of the phase diagram can be expanded by introducing anisotropic interactions, or by working with mixtures of particles. Under some conditions, as the strength of attraction or particle concentration is increased, suspensions fall out of equilibrium and glasses or gels are observed. Here we review links that have been recently made between experimental observations and statistical mechanics of suspension phase behavior and suggest where advances will be made in the future.

BACKGROUND

Colloidal suspensions display phase behavior similar to that of molecular systems.[1] Order/disorder transitions are referred to as fluid/crystal transitions. Suspension phases separate out into dilute and concentrated phases in equilibrium with each other, thus displaying gas/fluid transitions. Dense suspensions form glasses. Although in single-component molecular systems, the thermodynamic variables used to describe phase transitions are density, temperature, and pressure, in colloidal systems, the variables are density, temperature, osmotic pressure, and strength of interaction. The last variable distinguishes colloidal suspensions most strongly from molecular systems. For example, because the strength of interaction can be tuned by solution properties, at fixed temperature, one can map out an entire phase diagram.[1]

The connection between descriptions of phase transitions in molecular systems and colloidal systems is based on the use of the effective potential where solvent-mediated interactions are ascribed to the particle interaction energy. The basis of this mapping goes back to McMillan and Mayer,[2] and the treatment of colloidal suspensions containing one type of particle as a pseudo-one-component system has been extensively tested. As the range and strength of attractions and repulsions can be varied over an enormous range with colloidal particles, suspensions of nanoparticles offer a unique test bed for treatments of molecular phase behavior.

The phase behavior of repulsive systems has been extensively studied.[1,3] For monotonic repulsive interactions, only order/disorder and glass transitions are observed. The effects of attractions have seen less investigations because of the limited number of experimental systems where the strength of attraction can be varied systematically over a range around the average thermal energy in the system. The limited number of model systems where weak attractions can be modulated arises from the ubiquitous nature of van der Waals interactions and the large magnitudes of these attractions at contact for particles half a micron or larger in size. Realization that the strength of van der Waals forces at contact is limited by the granularity of matter (which limits the distance of closest approach) leads one to conclude that nanoparticles will experience only weak attractions. Indeed, this realization has led to startling connections between the phase behavior of nanoparticles or large molecules, and the simulations and analytical models for the phase behavior of particles experiencing weak, short-range attractions.[4–6]

Suspensions containing nanoparticles are widely used in the production of ceramics, consumer products, paints and inks, and quantum dots for optical applications. Nanoparticles are used because of their optical properties, the fact that the electronic states of nanoparticles are altered from their states in the bulk, and because they offer unique chemical and physical properties to the suspensions. Of growing interest in the assembly of nanoparticles are the factors that control the states into which nanoparticles assemble. Key examples can be found in biochemistry, where biological systems are controlled by the interactions and assembly of nanoparticles whereas the structure of globular macromolecules is often determined by X-ray diffraction derived from crystals. Thus advances in the use of nanoparticles and an understanding of complex biological systems will be tied to a greater understanding of particle interactions and how these alter the state of nanoparticle aggregation.

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120022897

The fact that many systems of interest interact weakly where the strength of attractions lies on the order of a few times the average thermal energy in the system means that many nanoparticle systems sample an ensemble of states and are subject to undergoing phase transitions. Progress toward unraveling the effects of particle shape, anisotropy of interaction, and details of the spatial extent of the particle interaction as well as the kinetics of the self-assembly process is being made through investigations of systems containing nanoparticle spheres with centrosymmetric interactions. In this article, we focus attention on these advances while returning to more complex systems at the end as an indication of the opportunities for advancement that lead into the future.

As discussed above, a key feature of suspensions as opposed to pure molecular systems lies in the fact that particle interactions can be tuned based on the composition of the continuous phase. Thus, as a result, small changes in solution variables such as pH, ionic type or composition, and polymer concentration or molecular weight or temperature can give rise to dramatic changes in the state of particle aggregation and the mechanical properties of the suspension.[1,3] As a specific example, suspensions of silica particles stabilized with octadecyl chains grafted to their surface and suspended in a near-index matching fluid can be crystallized, gelled, made to separate into two liquid phases, or separated into three equilibrium phases, depending on the amount and size of added nonadsorbing polymer.[7,8] The literature abounds in the number of examples in which interactions between various kinds of particles are manipulated to place them into useful structures. Although the details of how these structures are achieved will vary with the chemistry of the solid and fluid phases of interest, the underlying physical chemistry of the colloidal state will be common between different materials, thus offering general guidelines on how to achieve the desired properties.

The complexity of describing the interactions and phase behavior of colloidal suspensions is simplified to a certain extent by treating the suspension as an effective one-component system with the interactions between the particles mediated by the solvent and various additives. This allows the extensive literature developed in the field of molecular thermodynamics to be harvested to advantage. Of key concern to these systems is the particle interaction energy. Here we will focus on particles with hard cores (i.e., the particles cannot interpenetrate) that experience attractions or repulsions.

HARD SPHERE BEHAVIOR

Hard spheres represent a model system that contains much of the essential physics for describing behavior in dense systems and has been extensively studied in the thermodynamics and statistical mechanics literature.[1] The development of suspensions where particles interact very nearly as hard spheres has advanced the understanding of both fields. A significant advancement was made when links between the order/disorder phase transition in charged stabilized latex particles and computer simulations of the interactions of hard spheres were established. Since this initial connection, advances have been made in understanding the microstructure, thermodynamics, and nonequilibrium behavior of dilute and dense systems of hard spheres.

Hard spheres interact with each other only when they touch with an infinite repulsion at contact representing their impenetrable physical volume. Thus entropy changes because of different packings control suspension equilibrium microstructures. Fig. 1 is the plot of the experimentally measured static structure factor for hard sphere suspensions at different volume fractions.[9] The structure factor is an equilibrium property that is indicative of the average spacing between the hard sphere particles and can be determined experimentally by scattering techniques. At a volume fraction (ϕ) of 0.495, suspensions of Brownian hard spheres arrange in an ordered lattice such that their

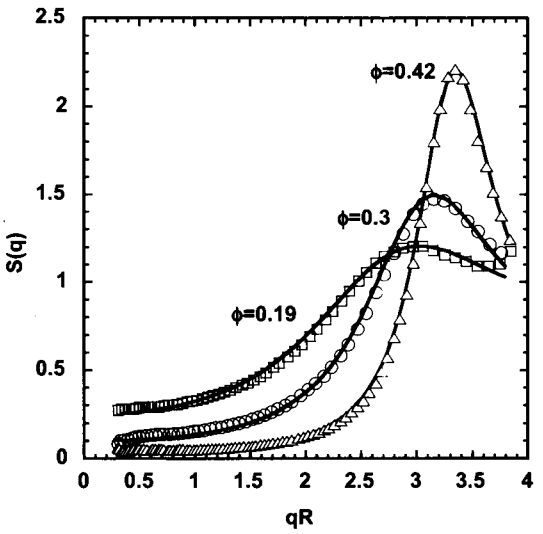


Fig. 1 Static structure $S(q)$ of hard sphere colloidal suspensions plotted as a function of the dimensionless wave vector qR. Here R is the sphere radius and q is the scattering vector [$q=4\pi/\lambda_1 * \sin(\theta/2)$, where λ_1 is the wavelength of incident radiation and θ is the scattering angle]. The solid lines are Percus–Yevick calculations of $S(q)$ for hard spheres. The open symbols are the experimental points as measured by small-angle X-ray scattering on the octadecyl-coated hard sphere silica system. The height of the first peak in the structure factor is indicative of the coherence of the first cage of particles and increases with increasing volume fraction. (From Ref. [9].)

entropy is maximized. The onset of freezing produces a solid at a volume fraction of 0.545 and the crystal can be further compressed to the closest packing density of 0.74. This phenomenon was first discovered through computer simulations and these predictions were first experimentally tested with colloidal suspensions.[1,10] Computer simulations demonstrate that face-centered cubic (FCC) microstructures have the lowest energy, but ground-based experiments and those done under microgravity conditions[11] (to reduce the effects of density mismatches) indicate that the particles tend to crystallize as randomly stacked hexagonal close-packed (HCP) structures, although an FCC microstructure slowly emerges if the crystals are allowed to age.

Nucleation rates, crystal growth velocities, and induction times for crystallization in hard sphere suspensions have been the subject of numerous computer simulations and experimental studies (commonly used techniques include light scattering and microscopy). Although the equilibrium states of the suspensions are not impacted by the particles being suspended in a dissipative continuous phase, all rate processes are influenced to greater or lesser degrees by this continuous phase. Accounting for these hydrodynamic interactions is essential in understanding the mechanical behavior in the low to intermediate volume fraction regime.

At elevated volume fractions approaching the glass transition ($\phi \sim 0.58$), suspension mechanics appears dominated by long-time self-diffusivities, which, once scaled on the infinite solution diffusivity, begin to be well described by models dominated by cage dynamics and cooperative rearrangements where hydrodynamic interactions are less important.[12] The approach of the glass transition results in the long-time self-diffusivity dropping to zero and a divergence in the zero shear rate suspension viscosity. Thus a complete description of suspension dynamic behavior requires a transition from regions where suspension mechanics is dominated by hydrodynamic interactions at low ϕ to one where cage dynamics dominates at elevated ϕ.[13]

Studies on the crystallization of hard spheres show a characteristic maximum in crystal nucleation rate as volume fraction is raised above 0.495 (Fig. 2).[14] To understand these observations, two different models have been developed: 1) the application of classical nucleation theory, which employs the combination of thermodynamic and kinetic descriptions that account for changes in self-diffusivity with volume fraction;[15] and 2) a completely kinetic description of nucleation,[16] which differs significantly from classical theories by making no reference to the solid–liquid surface tension. Both of the approaches predict a maximum in nucleation rate at $\phi \sim 0.56$, as measured by experiments. In the classical model, just above $\phi = 0.495$, the thermodynamic driving

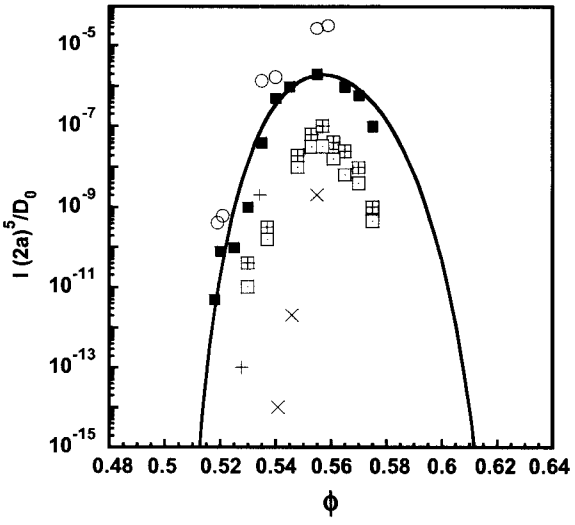


Fig. 2 Comparisons of predicted (solid line) and literature values of hard sphere nucleation rates as a function of particle volume fraction. (From Ref. [14].) The calculations are based on the kinetic model of Dixit and Zukoski. A maximum in the nucleation rate is observed at a volume fraction of 0.56. This is attributed to the competing effects of reduced particle diffusion and the increase in the thermodynamic driving force at elevated volume fractions.

force increases rapidly with volume fraction. The flux of particles to a cluster surface is determined by self-diffusion. As self-diffusivities decrease rapidly as ϕ approaches the glass transition near 0.58, the rate flux of particles to the nucleus surface drops to zero with increasing volume fraction. The competition between these two effects results in a maximum in nucleation rate. In the kinetic model, the flux of particles to a cluster surface is controlled by gradient diffusion. Just above $\phi = 0.495$, the flux of particles to a cluster surface increases simply because the bulk concentration increases, whereas the concentration at the cluster surface changes slowly. However, as ϕ further increases, the concentration difference between the bulk and the surface decreases, with the nucleation rate passing through a maximum. In this model, the flux to the surface is governed by gradient diffusion that has a much weaker dependence on volume fraction than self-diffusivity. A difference in the predictions of these two models lies in the volume fraction dependence of the size of a critical nucleus. Although the kinetic model predicts a minimum in critical nucleus size with increasing volume fraction, the classical approach predicts a monotonic decrease in critical cluster size with increasing volume fraction. There are two experimental data sets reporting volume fraction dependence of the critical nucleus size. In one, the critical nucleus size monotonically decreases, whereas in the second, a minimum in critical nucleus size

is reported. Consequently, the differences in these descriptions of hard sphere crystal nucleation will require further experimental and modeling work.

SOFT REPULSIONS

The addition of "soft" or long-range repulsions to colloidal particles can overcome the attractive van der Waals attraction to render them stable against aggregation. The most common method of electrostatic stabilization is the adsorption of charged species or dissociation of bound surface groups. The range of electrostatic repulsion is characterized by the Debye length (κ^{-1}) and can be regulated by adding salt to the system.[1] The Debye length scales inversely with the square root of the ionic strength of the suspension. Charged particles exhibit an order/disorder phase transition (Fig. 3)[17] similar to those seen in hard sphere systems, but at quite low volume fractions. The phase transition can be achieved by two means: by increasing the particle volume fraction, or by increasing the range of repulsion (decreasing the ionic strength). The FCC-ordered phase prevails at higher concentrations, or with shorter-range repulsions. As the range of repulsion is increased (say, by the removal of screening species in solution, thereby increasing the Debye length), a lower-density, body-centered cubic (BCC) packing forms the lowest free-energy ordered phase. BCC and FCC phases

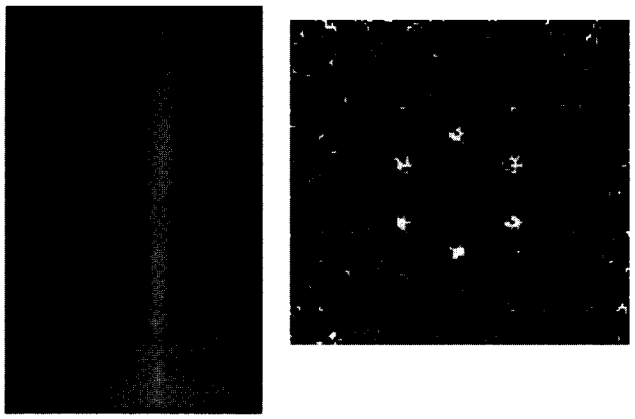


Fig. 3 Optical visualization of 229-nm (diameter) latex particle suspensions at a volume fraction of 0.53 at rest in a transparent shear cell and the resulting small-angle neutron scattering result. (From Ref. [17].) The observed iridescence is because of the crystalline packing of the latex particles. The small-angle neutron scattering pattern for the above suspension is given in the form of an isointensity plot where the highest intensity is represented by the lightest color. The spots are indicative of a microstructure with long-range orientational order consisting of HCP planes. (*View this art in color at www.dekker.com.*)

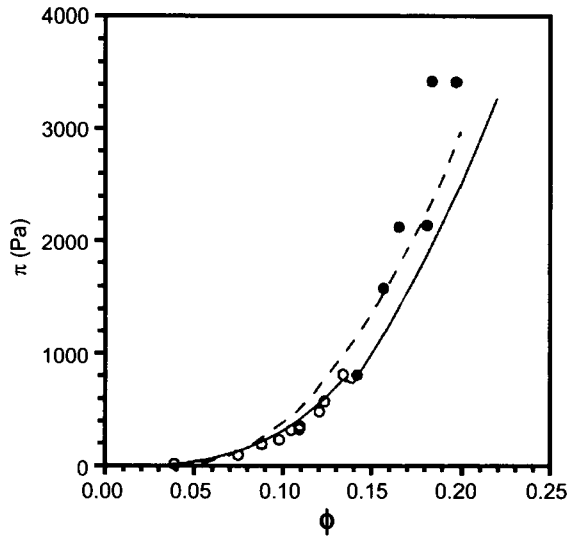


Fig. 4 Osmotic pressure as a function of volume fraction for suspensions of charged 35-nm (diameter) latex particles. The open circles represent the liquid phase (or the disordered state), whereas the solid phase of the suspension is shown in closed circles. The solid line is the prediction of perturbation theory and the dashed line is that of the cell model. The latex particles are highly charged and undergo a fluid/crystal transition at a low volume fraction of 0.14. (From Ref. [17].)

coexist under certain conditions in charged colloids as observed by the experiments of Chaikin et al.[18] The BCC–FCC–liquid triple point occurs when the interparticle spacing is 4.9 times the Debye length.

The equilibrium thermodynamics of the phase transitions in electrostatically stabilized suspensions is well captured by treating the particles as effective hard spheres where the effective hard sphere diameter is determined by the range and strength of the repulsion through self-consistent ways. More sophisticated approaches use one-component plasma and mean spherical approximation solutions of the pair distribution function or perturbation theory predictions of the suspension osmotic pressure and location of the phase boundaries (Fig. 4).[17,19,20] These models agree well with experimental observations when the particle surface charge is taken as the adjustable parameter. Good agreement is found for osmotic pressure through the disorder/order phase transition using these models.[18–21]

In very low ionic strength media, the ions that have dissociated from the particles provide a significant addition to screening of the electrostatic repulsions. Consequently, over extremely wide ranges of volume fractions, suspensions can sit on the edge of order/disorder transitions. Each time the particle concentration is increased, the resulting addition of counterions decreases the range of the interparticle repulsion such that the effective volume

fraction lies below 0.495 and the ordered phase remains unstable relative to the liquid phase.[22,23]

Encouraged by the accuracy of the effective hard sphere models in capturing the location of the phase boundary, attempts have been made to describe the kinetics of crystallization in charged sphere systems. A one-to-one mapping is complicated by different interpretations of how the diffusivity of the particle changes with the strength of repulsion. Recent experiments, simulations, and modeling efforts reveal the weaknesses of this approach. The location of the maximum in the nucleation rate, which occurs at a ϕ of 0.56 in hard spheres, depends sensitively on the strength and range of repulsions and is often not observed experimentally. In addition, depending on the strength and range of repulsions, drastically different nucleation rates are predicted at similar supersaturations, suggesting that a simple mapping of particle charge and Debye parameter into an effective particle size does not capture the subtleties of the kinetics of phase changes in systems with soft repulsions.[24,25]

WEAKLY ATTRACTIVE SUSPENSIONS

Investigations of the phase behavior of weakly attractive suspensions have seen relatively less work due primarily to a lack of model experimental systems, and/or methods independently characterizing the strength and range of the attractive forces. In the classical colloids literature, the goal has long been to control the state of aggregation that results when van der Waals attractions overwhelm repulsive forces. Van der Waals attractive potentials have contact values of $\sim AR/12\delta$, where $A \sim 1\text{–}10kT$ is the material-dependent Hamaker coefficient, R is the sphere radius, kT is the product of Boltzmann's constant and the absolute temperature, and δ is the distance of closest separation. Values for δ will depend on the material used and the morphology of the particles. However, several studies have suggested that $\delta \sim 0.65$ nm represents a minimum for liquid and molecularly smooth surfaces.[26] The study of the phase behavior of attractive colloidal systems has been limited by the large magnitude of this contact force when R is larger than ~ 50 nm. Up to a few years ago, most studies of attractive colloidal systems were based on understanding the rates of irreversible aggregation and the morphologies of the resulting aggregates.

Progress in controlling the magnitude of the attractive forces has been made by driving A to very small values by matching the index of refraction of the particle to that of the solvent, but even here, the number of model systems remains very small. More recently, there has been the realization that when R is small, the van der Waals attractions are small and the strength of attraction can be modulated in the range of a few kT.[6] Under these conditions, aggregation is reversible and suspensions can undergo phase transitions. This opens the door for connecting discussions of the thermodynamics of solutions and solute/solvent interactions with the language of suspensions, not surprisingly demonstrating the smooth transition in physical behavior as particle size grows from the molecular to the colloidal.

The two best studied systems of attractive colloidal suspensions are based on suppressing van der Waals attractions by index matching the particles while further suppressing attractions by coating the particles with short hydrocarbon chains that act as steric stabilizers. Silica particles coated with octadecyl hydrocarbon chains or trimethoxysilane materials, and polymethylmethacrylate particles stabilized by a layer of polyhydroxysteric acid suspended in a number of solvents have been used as model hard sphere systems. Attractions are generated in one of three ways. The first two involve changing the state of the coating/solvent interactions by altering the composition of the continuous phase, or by changing suspension temperature. Both methods drive the unfavorable stabilizing hair/solvent interactions, resulting in attractive interactions.

The third method of altering the strength of attraction involves addition of a nonadsorbing polymer. In a solution containing a nonadsorbing polymer, when two colloidal particles approach each other, the polymer is excluded from the gap between the particles. This exclusion leads to an imbalance in osmotic pressure of the polymer solution between the gap and in bulk solution. The imbalance drives the particles together and the resulting attraction is called "depletion" attraction. The strength of attraction is proportional to the polymer concentration, and the range of attraction is determined by the size ratio of polymer to colloid.[27,28] Depletion systems have the added advantage of providing a means of independently controlling both the strength and the range of attraction. The disadvantage of using depletion interactions as models of weakly attractive colloidal suspensions is that the suspension becomes a three-component system and the application of standard pseudo-one-component interaction potentials fails to describe the suspension thermodynamic behavior. Nevertheless, a great deal has been learned through experimental and modeling efforts focusing on depletion systems.

A third system for studying the phase behavior of weakly attractive suspensions involves native nanoparticles. As mentioned above, for these systems, van der Waals attractions are sufficiently small that suspensions become thermodynamically stable without steric coatings or index matching. The strength of the attraction can be modulated through changes in temperature, pH, ionic strength, or concentration of the nonadsorbing polymer.

For all of these systems, one of the keys to understanding phase behavior lies in developing methods to characterize the strength and range of the attractive interaction. One of the advantages of the depletion system is that analytical models are available for characterizing the strength of attraction in the limit of dilute particles. As the particle volume fraction increases, these models fail quantitatively and, in some cases, qualitatively.[7] However, the effects of the control variables (i.e., the concentration and molecular weight of the polymer) are captured in these models. Unfortunately, for the development of model systems, there are no a priori methods of knowing how changes in control parameters alter the strength or range of attraction in the octadecyl silica system as solvent composition or temperature is altered, or in nanoparticle suspensions with changes in continuous phase composition. However, it is the sensitivity of the interactions to the chemistry of the particles and solvent that offers the richness of the observed behavior and technological applications of many colloidal systems. As a result, there is growing attention being placed on developing methods for characterizing the effects of continuous phase composition on the strength of attraction of nanoparticles.

Attractive systems are expected to show disorder/order as well as gas/liquid phase transitions. In the colloidal field, the first models to capture this were those on depletion systems as described by Gast et al.,[28] where they demonstrated that when the range of the attraction was on the order of $0.3R$ or larger, a phase diagram similar to that expected for molecules would be observed. This phase diagram in a space where the inverse of the dimensionless strength of attraction is plotted as a function of particle concentration has a critical point and a triple point expected for systems showing liquid/liquid and liquid/solid phases. As the range of the attraction is reduced, the critical point drops below the triple point and there are only two stable phases predicted: fluids and crystals. Despite having a finite size potential and an attractive potential, thermodynamically stable gas/liquid transition is lost. The critical point lies below the fluid/solid phase boundary. The spinodal for this fluid/fluid phase boundary can be probed when the system is quenched rapidly, at which point the system spontaneously decomposes into a dilute and a concentrated colloidal suspension. A common interaction potential used to describe attractions is the Yukawa potential. The interaction energy $u(r)$ for Yukawa systems is given by:

$$u(r) = \begin{cases} \infty & r < 2R \\ \dfrac{-\varepsilon}{kT}\dfrac{e^{(R/\lambda)(1 - r/2R)}}{r/2R} & r \geq 2R \end{cases}$$

where λ/R is the range of attraction and ε/kT is the strength of attraction. The phase diagram of Yukawa fluids is given in Fig. 5[29] for different values of λ/R and, as mentioned above, the critical point begins to drop below the fluid/solid phase boundary for $\lambda/R=0.29$.

Understanding of the influence of the extent of the attractive well on phase behavior was advanced by understanding the phase behavior of fullerenes (C_{60}),[4] where detailed calculations showed that no liquid phase would be stable, and protein suspensions, where the existence of a metastable fluid/fluid phase transition was revealed.[30,31]

In the limit of very-short-range attractions between particles, the simplest model for the interparticle interaction is the adhesive hard sphere (AHS) potential.[1] The AHS potential is a two-parameter potential with Γ_{AHS} characterizing the strength of interactions and R the particle radius. Extensive simulations and density functional theory calculations have identified the phase boundary associated with the liquid/liquid transition, percolation, and liquid/solid transition. However, the AHS system

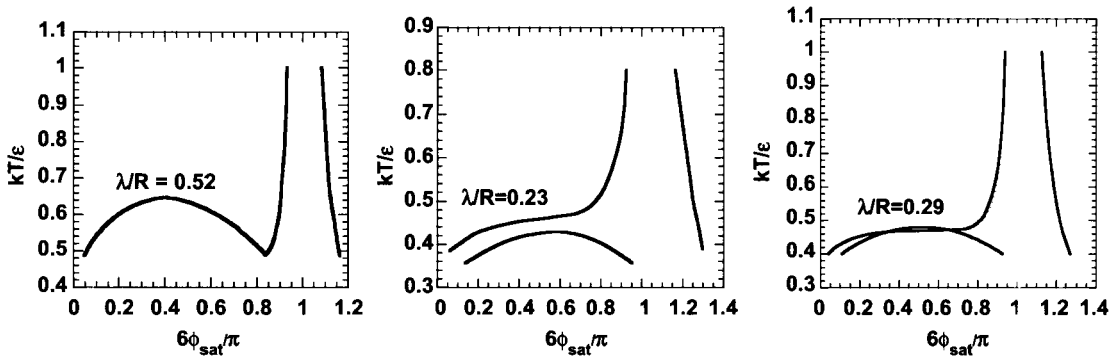


Fig. 5 Phase diagram calculations of the Yukawa potential using the perturbation theory for different ranges of attraction λ/R. (From Ref. [29].) When $\lambda/R<0.29$ (small ranges of attraction), the critical point begins to move below the fluid/solid phase boundary. The phase diagram of $\lambda/R=0.23$ is what is typically seen in a number of protein solutions and in solutions of inorganic nanoparticles.

contains certain limitations, and predictions of the phase boundary for all values of Γ_{AHS} are not possible.[32] However, as the range of attraction is decreased, the AHS, square well, and Yukawa interactions begin to converge if normalized by a common measure of the strength of attraction.[6] The general phase diagram resembles that of a very-short-ranged square well fluid or Yukawa fluid, with the liquid/liquid phase transition lying below the liquid/solid phase transition. Experimentally, this is seen in a number of protein solutions and in solutions of inorganic nanoparticles.

The utility of the square well interaction models in the limit of very narrow wells (i.e., approaching the AHS potential) for describing the thermodynamics of nanoparticle suspensions (proteins and inorganic silicotungstic acid) and in locating crystallization conditions has been repeatedly demonstrated.[6,23] Because of the inherent complexity in characterizing all the interaction forces (electrostatic, van der Waals, depletion, solvation and structural forces, hydrogen bonding, and salt bridges) between two nanoparticles in suspension, indirect measures of the overall interactions must be used. The experimental variables suggested for characterizing the interactions have been the inverse osmotic compressibility $(d\Pi/d\rho)_{sat}$ at the solubility point, the second virial coefficient (B_2), the relative solution viscosity (η/η_c), and the solubility ϕ_{sat}. Various techniques have been developed for extracting these parameters for model-free representations of phase behavior that become important for testing different interaction potentials. These studies demonstrated the universal nature of the solubility curve when the measured second virial coefficient ($B_2/16\pi R^3/3$) characterizing pair interactions was plotted against the solubility ϕ_{sat} (Fig. 6).[23] Agreement between theory and experiment is excellent. These results suggest that the crystallization boundaries of nanoparticle suspensions are relatively insensitive to the details of the interaction potential and that if two suspensions have the same $B_2/(16\pi R^3/3)$, then they will have the same solubility. These results have applications in the location of solution conditions giving rise to protein crystals.

Rates of nanoparticle crystallization have been the subject of controversy. In the first place, these crystals often nucleate slowly (Fig. 7).[33] Without this being the case, it would not be possible to probe the existence of the metastable liquid/liquid spinodal. Numerous studies have applied classical nucleation theory to extract solid/liquid surface tensions that are in broad agreement with experiments. Kulkarni and Zukoski[34] demonstrated that these surface tensions were in good agreement with predictions based on pair interaction models that are capable of predicting the location of the solubility boundary. The influence of the location of metastable liquid/liquid spinodal relative to the solubility curve on the rate of nucleation

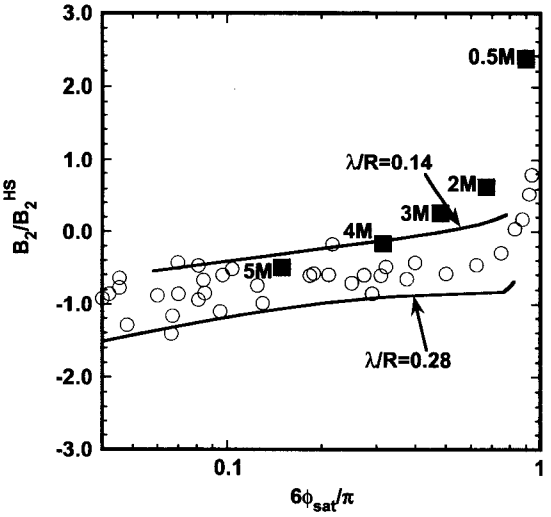


Fig. 6 Measured second virial coefficients of STA (solid squares) in different background salt concentrations compared with data on a number of proteins (lysozyme; Bovine Pancreatic Trypsin Inhibitor (BPTI) open circles) in different buffer solutions. The second virial coefficients are nondimensionalized with the hard sphere value and plotted against the solubility (volume fraction ϕ_{sat}) of the respective species. The solid lines are calculations of the attractive Yukawa potential with two different ranges of attractions λ/R of 0.14 and 0.28. The values of 0.14 and 0.28 indicate that attractions between the particles are short-ranged. The experimental data for silicotungstic acid (STA) (at high salt concentrations) and proteins collapse within the narrow range of attractions, which are only a fraction of the particle diameter. The collapse also indicates that proteins and STA are thermodynamically similar; if two suspensions have the same B_2, then they have the same solubility. This plot also provides an opportunity to extract interaction potential parameters for a given experimental system in a model-independent manner. (From Ref. [23].)

has been explored. Simulations and analytical models indicate that the enhanced compressibility of the suspension near the critical point will greatly enhance the rate of crystal nucleation.[35,36] Thus modeling efforts suggest that high-quality crystals can be grown near the critical point of the phase boundary. However, experimental verifications of these observations are difficult because of the high concentrations of proteins at the critical point ($\sim$400 mg/mL). There have been a number of experimental studies in measuring the nucleation rates of proteins under a variety of conditions.[37–39] Kinetic models of nucleation consistently overpredict experimental estimates. Dixit and Zukoski[36] attribute this to the use of simple centrosymmetric potentials in describing the kinetics of nucleation. More sophisticated patchy interactions that better approximate protein interactions with resulting modifications in the nucleation mechanism might provide an answer.

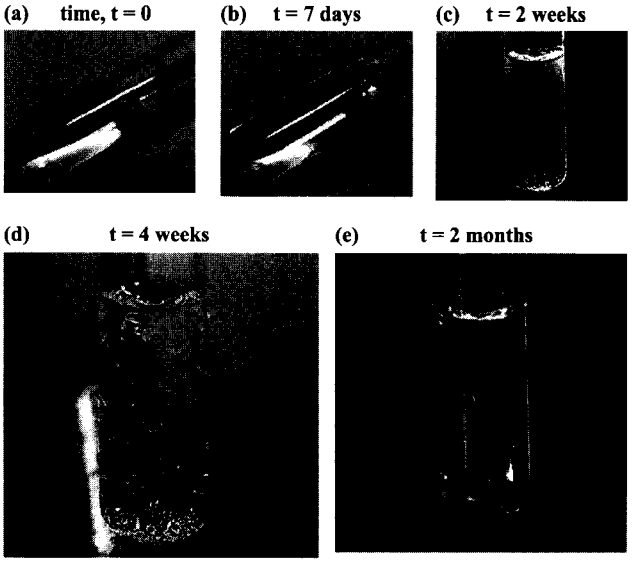


Fig. 7 Lysozyme solution in 0.6 M sodium phosphate buffer, pH 7.0, at an initial concentration of 100 mg/mL at 10°C. The solution starts fluidlike (a) and gels over a period of 1 week (b). Crystallites appear to nucleate within the gel (c), which grow (d) and finally form large ordered crystals (e), in equilibrium with a fluid of low protein concentration. A detailed study of the solubility and crystal nucleation kinetics revealed that these phenomena are largely controlled by the strength of attraction and that if two proteins have the same B_2, then they have the same kinetics of nucleation. (From Ref. [33].) (*View this art in color at www.dekker.com.*)

Connections between absolute rates of crystal nucleation and theoretical predictions have led to some controversies. Although many studies confirm that classical nucleation theory captures the supersaturation dependence on the nucleation rate, the absolute rate is very poorly described by these models. As with molecular systems, measuring the rate of nucleation and linking what can be measured to what is predicted by extant models remain at the heart of this controversy. A significant conclusion is that we have a poor understanding of what it is that controls the rates of crystal formation in supersaturated nanoparticle suspensions.[37–39]

As mentioned earlier, control over the strength and range of attractions can be achieved by the addition of nonadsorbing polymers to colloidal suspensions. The important variables in these systems are the polymer concentration (c_p), which determines the strength of attraction, and the size ratio of polymer to colloid (R_g/R), which determines the range of attraction. Recent experiments performed on silica[7,8] and polymethylmethacrylate model systems[27,40] characterize phase behavior as a function of particle volume fraction, R_g/R and c_p. In addition, osmotic compressibility and microstructures have been measured for the octadecylsilica suspensions suspended in decalin in the presence of polystyrene over a wide range of molecular weights.[8,9] A variety of phases are observed depending on the value of R_g/R and polymer concentration (Fig. 8).[8] Even though standard theories could predict the R_g/R value of 0.3 below which crystallization occurred, the predictions of solution miscibility with R_g/R at fixed colloid volume fraction are in direct contradiction to what is observed experimentally. Experimentally, solution miscibility improves as R_g/R increases at fixed ϕ, whereas standard theories predict the opposite trend. Comparisons with recently developed Polymer Reference Interaction Site Model (PRISM)[41,42] led to qualitative agreements with experimental data. PRISM is a three-component model (particles, polymer, and solvent) that treats the polymers as flexible chains of monomer units that have both conformational and translational entropy, which needs to be taken into account in describing

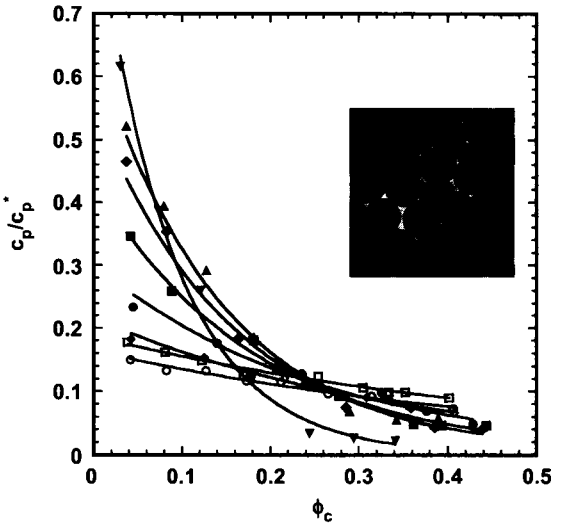


Fig. 8 Phase diagram of 100-nm (diameter) octadecanol-coated silica particles (inset figure) suspended in decalin in the presence of nonadsorbing polystyrene. (From Ref. [8].) The molecular weight of polystyrene is varied to achieve eight different size ratios of polymer to colloid (R_g/R) and the phase diagram determined for each R_g/R. Open symbols represent gel boundaries for R_g/R=0.025 (circles), 0.061 (squares), and 0.09 (diamonds). Filled symbols represent binodals for fluid–solid [R_g/R=0.243 (circles)] and fluid–fluid [R_g/R=0.525 (squares), 0.802 (diamonds), 1.10 (triangles), and 1.74 (reverse triangles)] phase separation. The solid lines (phase boundaries) are drawn to guide the eye. The colloidal silica suspensions are stable below the solid line. As more polymer is added and as one crosses the phase boundary, phase separation/gelation occurs depending on R_g/R.

the thermodynamics of colloid–polymer mixtures. The theory takes into account particle–particle, polymer–polymer, and polymer–particle interactions and qualitatively captures the right trends in phase behavior.

In addition to crystallization and liquid–liquid phase separation, suspensions experiencing very-short-range attractions are observed to gel when $R_g/R<0.1$.[7,8] Gelation is a phenomenon in which particles aggregate, forming a space-spanning structure that has a finite elastic modulus and yield stress. As mentioned before, gelation is a kinetic phenomenon and hence cannot be described by standard thermodynamic models. A dynamic model, which captures the arrest of particles with increasing polymer concentration and hence predicts the gelation boundary, is needed. The recently developed mode coupling theory (MCT), combined with PRISM (for colloid polymer mixtures), is capable of predicting gel boundaries.[43,44] MCT views the gel as particles trapped by a network of bonds that hinders the particle motion, resulting in a nonergodic state. Thus gelation (according to MCT) is caused by formation of long-lived cages. Gelation occurs at the point where the long-range self-diffusivity drops to zero. MCT is capable of predicting density autocorrelation functions as are measured by dynamic light scattering and has been applied to hard sphere systems where it captures many of the details of the colloidal hard sphere glass transition. It has also been applied to attractive suspensions where the gel line and dependence on R_g/R have been explored and the theory qualitatively captures the observed trends, particularly for ϕ above 0.2.

MCT is also successful in predicting another interesting phenomenon observed when increasing amounts of polymer are added to glassy hard sphere suspensions (i.e., $\phi>0.58$). The experimental observation is that, on addition of polymer (or when turning on a weak attraction), particles, on average, move closer together, but the coherence of the first shell of particles is disrupted. The result is that particles that are trapped by their nearest neighbors in the glassy state are now free to diffuse and the glass melts. However, on adding more polymer, the coherence of the first shell of particles is reestablished, the ability of particles to diffuse freely is diminished, and, with a sufficiently large polymer concentration, the suspensions gel. MCT is able to explain the observed phenomena based on the "cage" effect.[44] Thus, by introducing attractions in the system, one can go from glasses to liquids to gels. Recent studies on the dynamics of these systems demonstrate that the internal dynamics of gels is different form those of glasses. In particular, in glasses, the root mean square (RMS) displacement scales on the particle diameter, whereas in the gels, the RMS displacement of the particles scales on the localization length that is determined by the extent of the interparticle attraction and the particle volume fraction.

Of particular interest is the suggestion that suspensions of particles experiencing extents of attraction less than about $0.1R$ will not display thermodynamically stable phase transitions. Instead, with increasing volume fraction or with increasing strength of attraction, these suspensions gel or form irreversible aggregates. Thus the fractal flocs formed in irreversible aggregation (such as when salt is added to dilute latex suspensions) may result from short-range attractions rather than from strong attractions. This phenomenon can be circumvented in some cases. For example, if uniform suspensions are made up below the gel line in depletion systems and the particles are allowed to settle, crystals will form even when $R_g/R<0.1$. This observation indicates that understanding of the dynamics of attractive colloidal suspensions remains a fruitful area of research.

CONCLUSION

Suspensions of nanoparticles display a wide range of phase behavior—glasses, gels, crystals, and even separation into two liquid phases. The key variables that control the phase behavior are the range and strength of particle interactions. These variables can, in turn, be tuned in a number of different ways: by adding salt and polymer, by changing temperature, etc. Progress has been made in understanding the physics behind the phase separation process by working with well-characterized suspensions where the particles are spherical and interact with centrosymmetric potentials. This work has provided a basis for understanding the effects of particle anisotropy of interactions on phase behavior and the kinetics of the assembly process. As we look into the future, the challenges lie in extending the concepts and tools developed on simple systems to systems that are more complex, including anisotropic (rods and biological systems such as DNA) particles, and to mixtures of different kinds of particles (e.g., rods and spheres). Questions that must be answered include how interactions in such complex systems can be tuned to make them assemble into useful structures. For example, most studies of order/disorder phase transitions have focused on systems that produce cubic crystals. However, proteins are known to crystallize into a wealth of space groups. Is it possible to engineer the space group into which particles assemble by designing particle shape or surface chemistry? Similar questions can be asked about gels: Is it possible to alter the nature of the particle interactions to create gels with the desired volume fraction and flow properties by engineering particle size and shape? Answering these questions requires advances

in descriptions of phase behavior and kinetics of phase transitions as well as new synthetic approaches that result in particles of controlled surface chemistry. As an indication of the opportunities that are developing, we conclude this section by giving below a few examples involving systems with more complex interactions that lead to interesting phase behavior.

Tohver et al.[45,46] recently described a new mechanism for stabilizing colloidal particles. They noticed that when small amounts of charged nanoparticles were added to a gelled colloidal microsphere suspension (gelation occurs because of the attractive van der Waals forces), the gel melts and the suspension flows. However, increasing the amount of the nanoparticles flocculates the suspension. The initial stabilization mechanism is attributed to the segregation of the nanospheres on the surface of the microsphere colloidal particles because of their highly repulsive coulombic interactions in solution. This haloing effect results in a dense cloud of nanoparticles around each large particle such that when two large particles approach, they feel repulsive osmotic forces similar to that developed by overlapping double layers. This reduces the long-range van der Waals interactions between the colloidal spheres, thereby stabilizing the suspension. Increasing the amount of the nanospheres increases depletion attractions between spheres, thereby flocculating them at higher concentrations.

Polyelectrolytes (DNA, F-actin, and viruses) are complex molecules (in size and shape) that have a charge distribution on their surface, which is quite different from the uniform charge distribution on spheres that we dealt with in "Weakly Attractive Suspensions." In solution, charged polyelectrolytes repel each other (DNA in water) in the presence of monovalent ions, which is what is expected because like charges always repel. However, in the presence of multivalent ions, like charge attractions have been observed in a number of polyelectrolyte systems. Butler et al.[47] studied polyelectrolyte condensation using anionic rodlike M13 virus and a series of "tunable" divalent cations with the aim of understanding the role of ion multivalence and geometry in the phase change process. Using the experimental system, they were able to construct a multivalent ion–polyelectrolyte phase diagram and developed an experimentally motivated criterion for like-charge attraction based on the ion valence, ion size, and the Gouy–Chapman length. The phase diagram defines regions between disordered rods and regions where the rods have condensed into an aligned structure. This phase behavior is of interest because it demonstrates the ability of multivalent ions to produce a lowest free-energy state for electrostatically repulsive rods and provides insights into novel methods for organizing anisotropic particles.

Adams et al.[48] describe a rich phase diagram for mixtures of colloidal rods (filamentous bacteriophage fd virus) and spheres (polystyrene latex, polyethylene oxide, and polyethylene glycol). The phases they observed include: bulk demixing into rod-rich and rod-poor phases, and microphase separation into a variety of morphologies. One microphase consists of layers of rods alternating with layers of spheres; in another microphase, the spheres reversibly assemble into columns, which in turn pack into a crystalline array. Using the concepts of depletion attractions, they were able to predict the majority of the experimentally observed phases. The spheres and rods were modeled as hard objects, and the work done on the phase behavior of just pure hard spheres and hard rods aided in the development of a theory for the mixtures. However, the simple theory is unable to predict the existence of the columnar phase—the description of which remains a theoretical challenge.

The above cases were chosen to exemplify opportunities that exist for manipulating the states of aggregation of colloidal suspensions. Clearly, by working with strengths of attraction and repulsion on the order of a few times the average thermal energy of the system and by manipulating the degree of anisotropy of particle interactions, a wide range of structures can be built. This field will continue to require attention as the assembly of nanostructures continues to be of scientific and technological significance.

ACKNOWLEDGMENTS

This work was supported by the Nanoscale Science and Engineering Initiative of the National Science Foundation under NSF Award number DMR-0117792.

REFERENCES

1. Russel, W.B.; Saville, D.A.; Schowalter, W.R. *Colloidal Dispersions*; Cambridge University Press: Cambridge, 1989.
2. McQuarrie, D.A. *Statistical Mechanics*; Harper Collins: New York, 1976.
3. Hunter, R.J. *Foundations of Colloid Science*; Clarendon Press: Oxford, 1987; Vol. 1.
4. Hagen, M.H.J.; Meijer, E.J.; Mooij, G.; Frenkel, D.; Lekkerkerker, H.N.W. Does C-60 have a liquid-phase? Nature **1993**, *365* (6445), 425–426.
5. Hagen, M.H.J.; Frenkel, D. Determination of phase-diagrams for the hard-core attractive Yukawa system. J. Chem. Phys. **1994**, *101* (5), 4093–4097.
6. Rosenbaum, D.; Zamora, P.C.; Zukoski, C.F. Phase

behavior of small attractive colloidal particles. Phys. Rev. Lett. **1996**, *76* (1), 150–153.
7. Ramakrishnan, S.; Fuchs, M.; Schweizer, K.S.; Zukoski, C.F. Entropy driven phase transitions in colloid–polymer suspensions: Tests of depletion theories. J. Chem. Phys. **2002**, *116* (5), 2201–2213.
8. Shah, S.A.; Chen, Y.L.; Schweizer, K.S.; Zukoski, C.F. Phase behavior and concentration fluctuations in suspensions of hard spheres and nearly ideal polymers. J. Chem. Phys. **2003**, *118* (7), 3350–3361.
9. Shah, S.A.; Chen, Y.L.; Ramakrishnan, S.; Schweizer, K.S.; Zukoski, C.F. Microstructure of dense colloid–polymer suspensions and gels. J. Phys., Condens. Matter **2003**, *15* (27), 4751–4778.
10. Hachisu, S.; Kobayashi, Y. Kirkwood–Alder transition in monodisperse latexes: II. Aqueous latexes at high electrolyte concentration. J. Colloid Interface Sci. **1974**, *46*, 470–476.
11. Zhu, J.X.; Li, M.; Rogers, R.; Meyer, W.; Ottewill, R.H.; Russell, W.B.; Chaikin, P.M. Crystallization of hard-sphere colloids in microgravity. Nature **1997**, *387* (6636), 883–885.
12. Pusey, P.N. Colloidal Suspensions. In *Liquids, Freezing and Glass Transition*; Hansen, J.P., Levesque, D., Zinn-Justin, J., Eds.; Elsevier: Amsterdam, 1991; 763–942. Part II.
13. Fuchs, M.; Cates, M.E. Theory of nonlinear rheology and yielding of dense colloidal suspensions. Phys. Rev. Lett. **2002**, *89* (24), 248304. (1–4).
14. Dixit, N.M.; Zukoski, C.F. Kinetics of crystallization in hard-sphere colloidal suspensions. Phys. Rev., E **2001**, *64* (041604), 1–10.
15. Russel, W.B. On the dynamics of the disorder order transition. Phase Transitions. **1990**, *21* (2–4), 127–137.
16. Dixit, N.M.; Zukoski, C.F. Kinetics of crystallization in hard-sphere colloidal suspensions. Phys. Rev., E **2001**, *6404* (4), 1–10.
17. Chen, L.B. *The Dynamic Properties of Concentrated Charge Stabilized Suspensions*; Ph.D. Thesis, University of Illinois at Urbana-Champaign: Urbana, 1991.
18. Chaikin, P.M.; Pincus, P.; Alexander, S.; Hone, D. Bcc–Fcc, melting and reentrant transitions in colloidal crystals. J. Colloid Interface Sci. **1982**, *89* (2), 555–562.
19. Voegtli, L.P.; Zukoski, C.F. Adsorption of ionic species to the surface of polystyrene latexes. J. Colloid Interface Sci. **1991**, *141* (1), 92–108.
20. Voegtli, L.P.; Zukoski, C.F. A perturbation treatment of the order–disorder phase-transition in colloidal suspensions. J. Colloid Interface Sci. **1991**, *141* (1), 79–91.
21. Hone, D.; Alexander, S.; Chaikin, P.M.; Pincus, P. The phase-diagram of charged colloidal suspensions. J. Chem. Phys. **1983**, *79* (3), 1474–1479.
22. Beresfordsmith, B.; Chan, D.Y.C. Highly asymmetric electrolytes—A model for strongly interacting colloidal systems. Chem. Phys. Lett. **1982**, *92* (5), 474–478.
23. Ramakrishnan, S.; Zukoski, C.F. Characterizing nanoparticle interactions: Linking models to experiments. J. Chem. Phys. **2000**, *113* (3), 1237–1249.
24. Dixit, N.M.; Zukoski, C.F. Pseudo-steady rates of crystal nucleation in suspensions of charged colloidal particles. J. Phys., Condens. Matter **2003**, *15* (10), 1531–1552.
25. Palberg, T. Crystallization kinetics of repulsive colloidal spheres. J. Phys., Condens. Matter **1999**, *11* (28), R323–R360.
26. Israelachvili, J.N. *Intermolecular and Surface Forces*; Academic Press: London, 1991.
27. Lekkerkerker, H.N.W.; Poon, W.C.K.; Pusey, P.N.; Stroobants, A.; Warren, P.B. Phase-behavior of colloid plus polymer mixtures. Europhys. Lett. **1992**, *20* (6), 559–564.
28. Gast, A.P.; Hall, C.K.; Russel, W.B. Phase separations induced in aqueous colloidal suspensions by dissolved polymer. Faraday Discuss. **1983**, *76*, 189–201.
29. Ramakrishnan, S. *Thermodynamics and Phase Behavior of Nanoparticles*; M.S. Thesis, University of Illinois at Urbana-Champaign: Urbana, 1998.
30. Lomakin, A.; Asherie, N.; Benedek, G.B. Monte Carlo study of phase separation in aqueous protein solutions. J. Chem. Phys. **1996**, *104* (4), 1646–1656.
31. Asherie, N.; Lomakin, A.; Benedek, G.B. Phase diagram of colloidal solutions. Phys. Rev. Lett. **1996**, *77* (23), 4832–4835.
32. Stell, G. Sticky spheres and related systems. J. Stat. Phys. **1991**, *63* (5–6), 1203–1220.
33. Kulkarni, A.M. *An Investigation of Phase Behavior and Crystal Nucleation Kinetics in Solution of Globular Protein*; Ph.D. Thesis, University of Illinois at Urbana-Champaign: Urbana, 2001.
34. Kulkarni, A.M.; Zukoski, C.F. Nanoparticle crystal nucleation: Influence of solution conditions. Langmuir **2002**, *18* (8), 3090–3099.
35. tenWolde, P.R.; Frenkel, D. Enhancement of protein crystal nucleation by critical density fluctuations. Science **1997**, *277* (5334), 1975–1978.
36. Dixit, N.M.; Zukoski, C.F. Crystal nucleation rates for particles experiencing short-range attractions: Applications to proteins. J. Colloid Interface Sci. **2000**, *228* (2), 359–371.
37. Dixit, N.M.; Kulkarni, A.M.; Zukoski, C.F. Comparison of experimental estimates and model

predictions of protein crystal nucleation rates. Colloids Surf., A Physicochem. Eng. Asp. **2001**, *190* (1), 47–60.
38. Galkin, O.; Vekilov, P.G. Direct determination of the nucleation rates of protein crystals. J. Phys. Chem., B **1999**, *103* (49), 10965–10971.
39. Galkin, O.; Vekilov, P.G. Control of protein crystal nucleation around the metastable liquid–liquid phase boundary. Proc. Natl. Acad. Sci. **2000**, *97* (12), 6277–6281.
40. Ilett, S.M.; Orrock, A.; Poon, W.C.K.; Pusey, P.N. Phase-behavior of a model colloid–polymer mixture. Phys. Rev., E **1995**, *51* (2), 1344–1352.
41. Fuchs, M.; Schweizer, K.S. Structure of colloid–polymer suspensions. J. Phys., Condens. Matter **2002**, *14* (12), R239–R269.
42. Fuchs, M.; Schweizer, K.S. Structure and thermodynamics of colloid–polymer mixtures: A macromolecular approach. Europhys. Lett. **2000**, *51* (6), 621–627.
43. Bergenholtz, J.; Fuchs, M. Gel transitions in colloidal suspensions. J. Phys., Condens. Matter **1999**, *11* (50), 10171–10182.
44. Bergenholtz, J.; Poon, W.C.K.; Fuchs, M. Gelation in model colloid–polymer mixtures. Langmuir **2003**, *19* (10), 4493–4503.
45. Tohver, V.; Smay, J.E.; Braem, A.; Braun, P.V.; Lewis, J.A. Nanoparticle halos: A new colloid stabilization mechanism. Proc. Natl. Acad. Sci. **2001**, *98* (16), 8950–8954.
46. Tohver, V.; Chan, A.; Sakurada, O.; Lewis, J.A. Nanoparticle engineering of complex fluid behavior. Langmuir **2001**, *17* (26), 8414–8421.
47. Butler, J.C.; Angelini, T.; Tang, J.X.; Wong, G.C.L. Ion multivalence and like-charge polyelectrolyte attraction. Phys. Rev. Lett. **2003**, *91* (2), 028301. (1–4).
48. Adams, M.; Dogic, Z.; Keller, S.L.; Fraden, S. Entropically driven microphase transitions in mixtures of colloidal rods and spheres. Nature **1998**, *393* (6683), 349–352.

Phase Transfer of Monosaccharides Through Noncovalent Interactions

P

Elizabeth K. Auty
Anthony P. Davis
University of Bristol, Bristol, United Kingdom

INTRODUCTION

Sugars constitute one of the major classes of biological building blocks. On one hand, they are important for physical construction (e.g., as cellulose) and energy storage (e.g., as starch). On another more subtle level, they are used by nature to label cells and biomolecules and thus to mediate a variety of biological events. Thus saccharide motifs are central to cell–cell recognition, infection of cells by pathogens, and many aspects of the immune response.[1–8]

Carbohydrates are hydrophilic molecules, which are usually freely soluble in water and show little inclination to enter organic media. Nonetheless, there are good reasons why one might wish to promote phase transfer of these species. First, there is a potential large-scale technological application in the production of "high-fructose syrup." Hydrolysis of starch and enzymatic treatment can be used to produce a mixture of fructose and glucose, which is ca. 42% of the former. Higher levels are desirable as fructose is exceptionally sweet-tasting, and there is much interest in transport processes that might be used in enrichment procedures.[9,10] Second, there are biomedical applications for carbohydrate phase transfer. The most immediate might be in the management of diabetes, where an effective, reversible extraction system could be used as the basis for a glucose sensor.[11–14] In addition, there is fundamental interest in saccharide transport across biological membranes and the possibility of delivering carbohydrate-like drugs using phase-transfer agents. Third, the study of carbohydrate recognition has become a major area of supramolecular chemistry,[15,16] and carbohydrate phase transfer has proved useful in this context. The key challenge is to bind carbohydrates in direct competition with liquid water (their natural environment). However, it is difficult to design and handle receptors that are fully soluble in water. Extraction experiments, involving aqueous and organic phases, provide a straightforward solution to the problem. The receptor, dissolved in the organic solvent, is equilibrated with an aqueous solution of the substrate. The amount of substrate transferred to the organic phase can be measured and used as a qualitative indication of receptor efficiency. Quantification is possible in principle,[17] although to date this has not been realized for carbohydrate substrates.

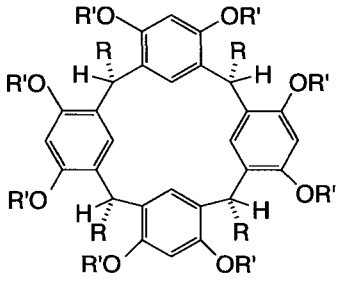
Broadly speaking, two approaches have been taken toward carbohydrate recognition in general, and toward carbohydrate phase transfer in particular. One is based on the reaction of boronic acids with diols to give cyclic boronate esters.[16,18] This strategy can be highly effective but, being based on covalent B—O bond formation, lacks biological relevance. The alternative relies on noncovalent interactions, and thus may be seen as "biomimetic."[15] This entry discusses systems belonging to the second category only, focusing on the transfer/transport of simple monosaccharide substrates into/across nonpolar phases. The account is divided into three sections covering: 1) preorganized receptor molecules; 2) less organized, micellar systems; and 3) self-assembling channels.

PHASE TRANSFER OF MONOSACCHARIDES BY PREORGANIZED RECEPTORS

Monosaccharides are relatively large as substrates for supramolecular chemistry, possessing irregular arrays of divergent functional groups. Being mostly hydroxyls, these substituents are capable of hydrogen bonding but do not form especially strong and specific interactions (unlike, for example, carboxylate or ammonium units). Thus carbohydrate recognition is a difficult undertaking, even in the absence of water. Competition from water molecules adds to the challenge, given that the receptor must distinguish the target from competitors bearing the same functional groups. Preorganization, the correct positioning of binding groups and surfaces, is the key to solving such problems.[19] Cyclization is a useful strategy for controlling the shape of large, potentially flexible frameworks, and it is not surprising that macrocycles have featured strongly in supramolecular chemistry.[20] Macrocycles composed of rigid units possess still fewer conformational options, so that the cyclophane architecture (a macrocycle incorporating one or more aromatic rings) is especially popular.[21,22] Most of the carbohydrate receptors with demonstrated phase transfer capability belong to this family of molecules.

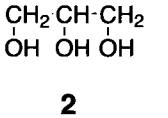
Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120018332

The first such system was the receptor **1a**, synthesized by Aoyama et al.[23,24] from resorcinol and dodecanal. This system is furnished with eight phenolic OH groups, capable of hydrogen bonding to the substrate hydroxyls, and four undecyl side chains, which maintain lipophilicity and organic solubility. **1a** was determined by ^1H nuclear magnetic resonance (NMR) to be the all-cis, bowl-shaped isomer shown, and did indeed prove soluble in apolar media such as benzene or CCl_4. Dissolved in these solvents, it was capable of carrying 4 Eq of water or glycerol **2** into the organic phase. When an extraction was performed using an equimolar mixture of water and **2** ([**2**]=[water]=11 M), only water was transferred into the organic medium with little coextraction of glycerol evident. However, when a 5.5-M solution of ribose **3** was used as the aqueous phase, the monosaccharide was extracted with water to form a complex of estimated stoichiometry: **1a**:**3**:H_2O=1:1:2. Interestingly, the ribose was found to be extracted exclusively in the α-pyranose form shown, in which all the OH groups are cis. This point is discussed further below. When the *O*-acetylated macrocycle **1b** was used in the extraction experiments, neither **2** nor **3** was transferred to the organic phase, suggestive of OH···OH hydrogen bonding as the key binding interaction.

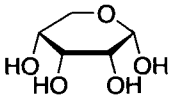


1a R = $(CH_2)_{10}CH_3$, R' = H

1b R = $(CH_2)_{10}CH_3$, R' = $COCH_3$



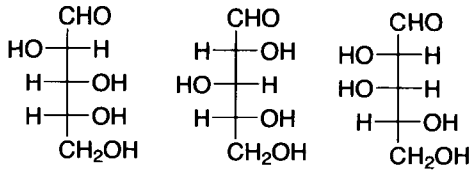
2



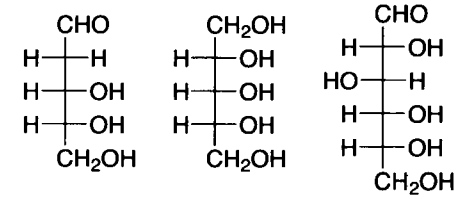
3

Aoyama et al. proceeded to study the extraction by **1a** of pentoses **4**–**8** and hexoses **9**–**13**. A summary of the results is given in Table 1.

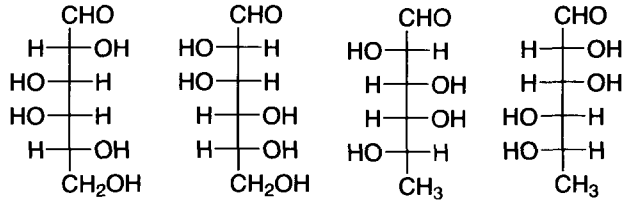
Two factors seem to dominate extractability by **1a**. The first, unsurprisingly, is the hydrophilicity of the sugar [e.g., D-galactose **10** is barely detected in the organic phase, whereas L-fucose (6-deoxy-L-galactose, **12**) is readily extracted]. The second is the relative configuration of the substrates at C_3 and C_4. Extraction seems to be favored by a cis arrangement in the Fischer projection (e.g., **3**, **4**, **7**, and **12**), which translates to cis in the pyranose form of the carbohydrates. The results support a face-to-face binding geometry in which OH groups emerging from the same face of the saccharide are best able to form multiple interactions with the receptor.



4 **5** **6**



7 **8** **9**



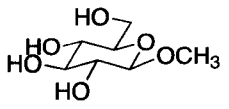
10 **11** **12** **13**

Later work showed that methyl-β-D-glucoside **14** could also be extracted by **1a** from water into CCl_4, despite the absence of *cis*-1,2-diol units.[25] However, in this case, a 2:1 receptor/substrate ratio suggested a "sandwich" structure, in which receptor molecules bind to both faces of the guest. The α-anomer **15**, for which the sandwich structure is probably not possible, was very poorly extracted under the same conditions.

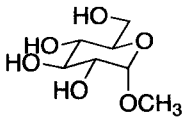
Table 1 Extraction of sugars from water into CCl_4 by **1a**[a]

Sugar	Sugar/**1a**[b]
D-ribose **3**	0.5
D-arabinose **4**	0.1
D-xylose **5**	vs[c]
D-lyxose **6**	vs[c]
2-deoxy-D-ribose **7**	0.8
Adonitol **8**	vs[c]
D-glucose **9**	vs[c]
D-galactose **10**	vs[c]
D-mannose **11**	vs[c]
L-fucose **12**	1.0
6-deoxy-L-mannose **13**	0.1

[a] $[Sugar]_{aq}$ = 2.4 M, $[\mathbf{1a}]_{org}=0.9\times10^{-2}$ M.
[b] Molar ratios sugar/**1a** appearing in the organic phase.
[c] Very small; ≤0.03.
Source: Ref. [24].



14



15

The second example of carbohydrate recognition by preorganized, macrocyclic receptors involved the steroid-derived "cholaphanes" **16**.[26,27] These molecules possess substantial cavities capable of fully surrounding a monosaccharide substrate. Up to six H-bonding groups (4 × OH and 2 × CONH) are available to interact with the carbohydrate, whereas the steroidal framework provides a lipophilic exterior compatible with organic media. A second generation of cholaphanes **17** featured externally directed alkyl groups to enhance lipophilicity. In this case, contraction of the side chain in steroidal precursors was used to vary cavity size and flexibility.[28]

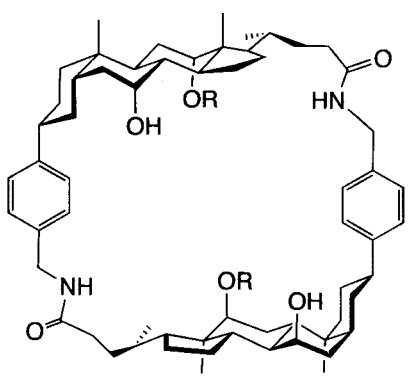
Receptors **16** and **17** were studied mainly by NMR in homogeneous $CDCl_3$ solution, employing organic-soluble glycosides such as **18** as substrates.[26–28] However, some extraction experiments were also performed by employing methyl-β-D-glucoside **14**.[28] The NMR data were compatible with 1:1 complex formation, and allowed the estimation of binding constants (K_a). For example, the following values were obtained for binding to **18**:**16**, $K_a = 3100\ M^{-1}$; **17a**, $K_a = 1560\ M^{-1}$; **17b**, $K_a = 600\ M^{-1}$; and **17c**, $K_a = 1305\ M^{-1}$. In the extraction experiments, solutions of the receptors in chloroform were stirred with aqueous solutions of **14** (1–to 2.5 M). Perhaps surprisingly, the order of effectiveness was **17a** > **17b** > **16** > **17c**, failing to parallel the affinities to **18**. The most efficient, **17a**, was capable of extracting ~0.2 Eq of **14** from a 1.75-M aqueous solution, and was also shown to transport this substrate through a chloroform barrier ("U-tube" experiment). Positive results were not obtained with the more hydrophilic glucose. Although not especially effective, this system performed a useful service in relating binding constants to extraction capabilities. Broadly speaking, a K_a of roughly $10^3\ M^{-1}$ in chloroform proved sufficient to extract a moderately hydrophilic carbohydrate (**14**) from fairly concentrated (~2 M) aqueous solutions.

A third family of macrocycles for carbohydrate phase transfer was developed by Inouye et al. Their initial system, exemplified by **19**, was especially targeted at ribofuranosides.[29] As shown, it was designed to provide a

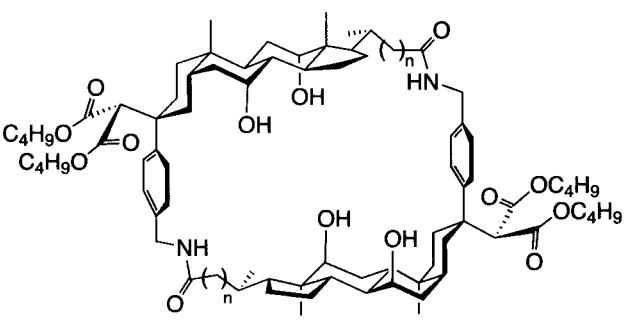
Table 2 Solubilization of sugars in chloroform by **19** and **22** ($R' = C_{11}H_{23}$)[a]

Sugar	Sugar/19[b,c]	Sugar/22 ($R' = C_{11}H_{23}$)[b,d]
D-ribose **3**	1.0	0.65
D-arabinose **4**	0.06	0.53
D-xylose **5**	0.5	0.52
D-lyxose **6**	0.63	0.66
2-deoxy-D-ribose **7**	0.7	1.0
D-glucose **9**	—	0.45
D-galactose **10**	—	0.32
D-mannose **11**	—	0.39
D-fructose **23**	0.2	0.52

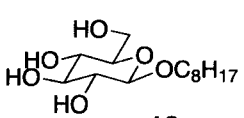
[a] $[Receptor]_{org} = 10^{-2}$ M.
[b] Molar ratios sugar/receptor appearing in solution.
[c] *Source*: Ref. [29].
[d] *Source*: Ref. [30].



16a R = H
16b R = CH_2Ph



17a n = 0
17b n = 1
17c n = 2



18

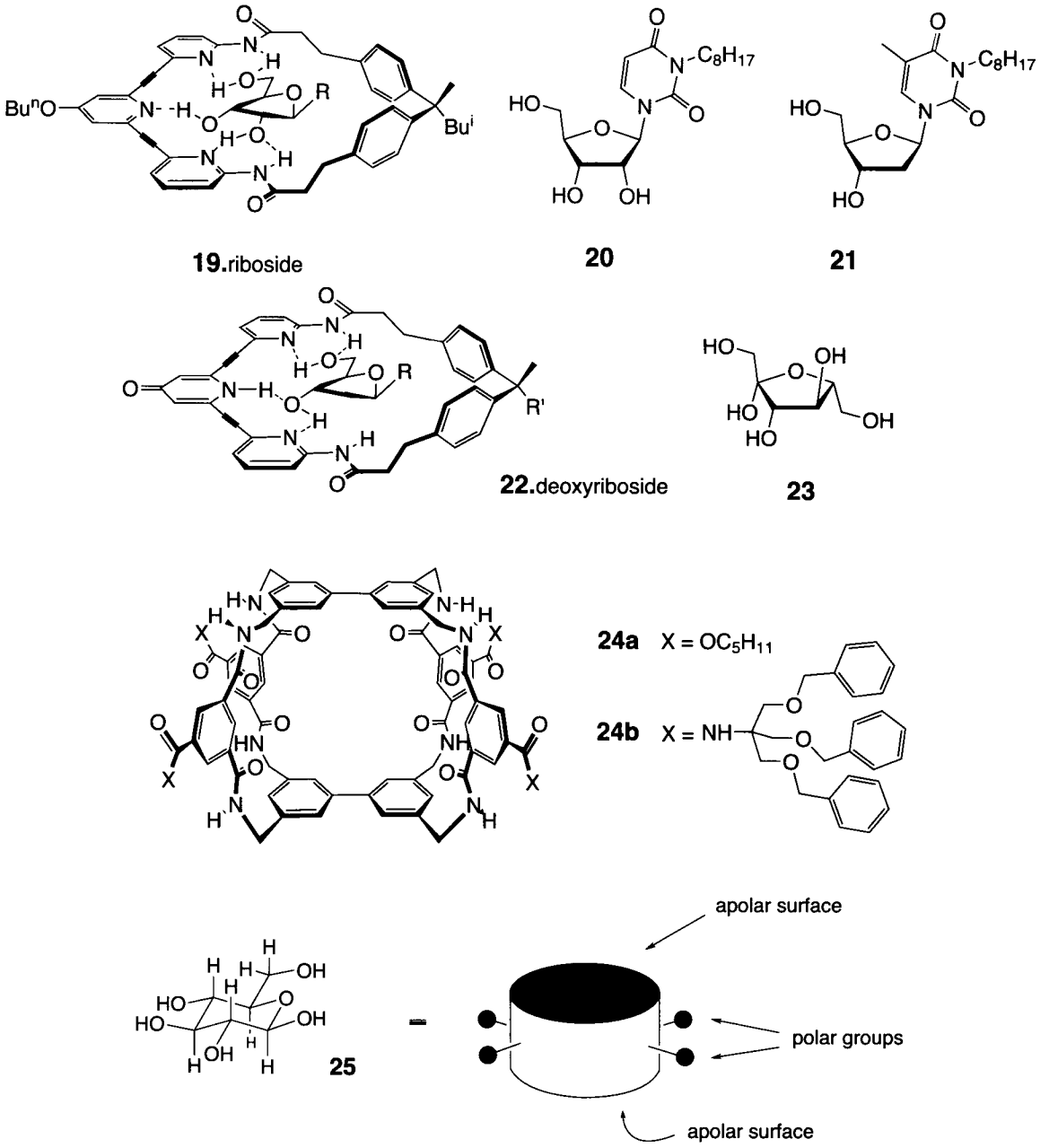
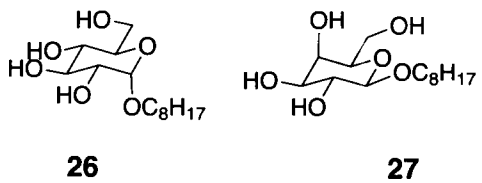


Fig. 1 D-glucose **9** in β-pyranose form **25**, interpreted as polar and apolar binding regions. (From Ref. [31].)

series of H-bonding interactions with the three hydroxyl groups of these substrates. In NMR titration experiments with the organosoluble β-D-ribofuranoside **20** and the 2-deoxy substrate **21**, binding constants of 10,000 and 690 M^{-1} were measured, respectively. The lower value for **21** is not surprising, given that it can form fewer hydrogen bonds with the receptor. The group then converted the central alkoxypyridine unit to a pyridone, as in **22**.[30] It was hoped that the change from H-bond acceptor to donor would make the receptor more compatible with



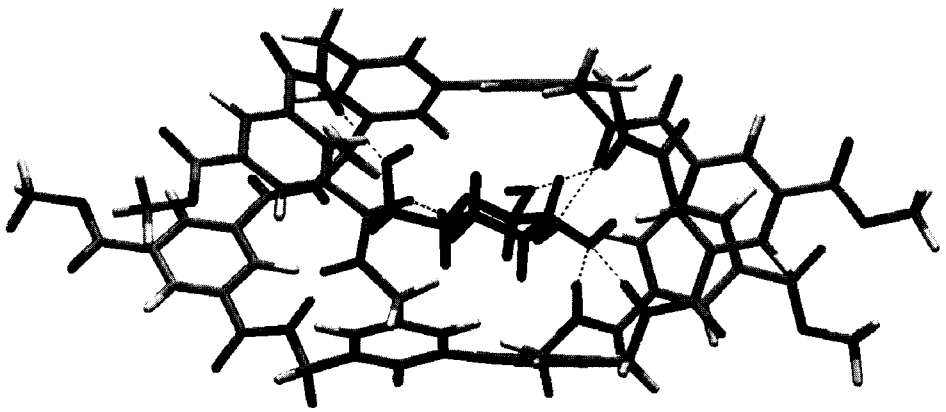


Fig. 2 Receptor **24**, modelled as the tetramethyl ester, binding D-glucose **9** as the β-pyranose. The configuration shown is the baseline conformation from a 1000-step Monte Carlo molecular mechanics study. Intermolecular hydrogen bonds are shown as broken lines. (From Ref. [31].)

2-deoxyfuranosides, as shown. Indeed, the binding constant of **22** (R = Bu^t) to **21** in $CDCl_3$ was found to be 19,000 M^{-1}.

The phase-transfer capabilities of **19** and **22** were explored through solubilization experiments employing solid carbohydrate substrates and chloroform as solvent. The results are summarized in Table 2. The data show that the affinities achieved by **19** and, especially, **22** are sufficient to solubilize a full range of monosaccharides (although not necessarily extract them from water). They also reveal genuine, receptor-based selectivity. To amplify this point, a series of figures (binding constants, extraction levels) recorded for a single receptor may not be especially informative. Differences may simply reflect the fact that some substrates are more easily bound (''stickier'') than others. However, when two receptors show distinctly different preferences, receptor structure is shown to be a major factor. In this case, **19** extracts D-ribose **3** more effectively than 2-deoxy-D-ribose **7** and fails with pyranoses **9–11**, whereas **22** prefers **7** to **3** and succeeds quite well with **9–11**.

Macrocyclic receptors **16**, **17**, **19**, and **22** are able to surround their substrates but not encapsulate them. A further advance was made by the macrotricyclic system **24**, which provides a preorganized, essentially enclosed binding site specifically designed for a monosaccharide.[31,32] In this case, the target was glucose **9**, in its native pyranose form. As shown in Fig. 1, the all-equatorial pattern of polar substituents in β-D-glucopyranose **25** allows the substrate to be viewed as a disk, with apolar surfaces top and bottom (corresponding to axial CH units) and a polar circumference. Accordingly, **24** consists of two parallel apolar units (biphenyls) connected by bridges containing H-bond donors/acceptors (amides). Modelling (Fig. 2) showed that **25** could fit in the cavity and make at least six intermolecular hydrogen bonds, as well as apolar CH–π contacts.[31] The structure is reminiscent of that found in carbohydrate-binding proteins, where aromatic surfaces are often found in contact with lipophilic patches on saccharide substrates.[15]

Receptor **24** was initially prepared as tetraester **24a** and studied in homogeneous solution. Octyl glycosides **18**, **26**, and **27** were used as substrates. Binding constants were obtained by NMR in $CDCl_3$-CD_3OH (92:8) and by fluorescence titration in $CHCl_3$. The results, summarized in Table 3, confirmed that the enclosed binding site yielded significant gains in affinity and selectivity. Thus the complex to β-glucoside **18** was formed with $K_a \sim 10^3\ M^{-1}$ even in the presence of 8% by volume of competing methanol, and in pure $CHCl_3$ (as used for previous systems), the value was raised to $3 \times 10^5\ M^{-1}$. In both media, α-glucoside **26** was far less strongly bound.

The phase-transfer properties of **24** were first tested by stirring **24a** with D-glucose **9** in $CDCl_3$.[31] Analysis by NMR indicated that ~0.9 Eq of **9** was solubilized, representing (as expected) an improvement on **22**. Remarkably, the glucose was detected as a 7:1 mixture of β-pyranose/α-pyranose, presumably reflecting the preference for β-glucosyl revealed in Table 3.[a]

For liquid–liquid extraction, it transpired that increased lipophilicity was required, so **24** was prepared as tetraamide **24b**.[32] Extraction experiments were performed in chloroform/water, with the results shown in Table 4. All the substrates were extracted from a 1-M aqueous solution including, for the first time, the hexoses **9–11**. Notable selectivity was shown for the ''all-equatorial'' saccharides glucose **9** and xylose **5**. Moreover, glucose, the specific

[a]The β:α ratio should normally be close to 1:1 in organic solvents.

Table 3 Binding constants (K_a) of octyl glycosides to macrotricycle **24a**

Substrate	K_a [M^{-1}] in $CDCl_3$-CD_3OH (92:8)[a]	K_a [M^{-1}] in $CDCl_3$-CD_3OH (92:8)[b]
β-D-glucopyranoside **18**	980 (±2%)	300,000 (±6%)
α-D-glucopyranoside **26**	20 (±23%)	13,000 (±8%)
β-D-galactopyranoside **27**	220 (±5%)	110,000 (±12%)

[a]^{1}H NMR titrations.
[b]Fluorescence titrations.
Source: Ref. [31].

target of **24**, was extracted from aqueous solutions as low as 0.1 M in concentration. Physiological glucose concentrations, at 0.005 M, are still out of reach, but with further development, this system could ultimately serve as the basis of a clinically relevant glucose sensor.

PHASE TRANSFER OF MONOSACCHARIDES BY LESS ORGANIZED SYSTEMS

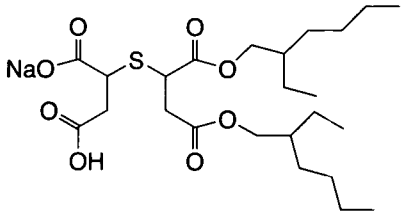
The essential characteristic required for carbohydrate phase transfer is an amphiphilic structure with a polar interior and an apolar exterior. Although preorganization of the polar interior should, in principle, be advantageous, it turns out that less organized structures can also be successful. One example is provided by thiosuccinate derivative **28**, studied by Greenspoon and Wachtel.[33] At a concentration of 50 mM, **28** was shown by x-ray scattering experiments to form "inverse micelles" (polar interior, apolar exterior) in a mixture of chloroform/cyclohexane, 1:1. Addition of D-glucose **9** (solid monohydrate) followed by centrifugation gave a clear solution, which was found by NMR to contain one molecule of glucose per nine molecules of **28**. It appears that the assembly of **28** can bind the glucose directly, without a solvation shell of water. When water was added, the glucose-derived NMR signals shifted, indicating a change in environment. Further NMR experiments with **29** as substrate, and electron spin resonance spectroscopy (ESR) studies employing **30**, confirmed that the monosaccharide unit was associated with the micelles.

Kobayashi et al.[34] studied **31**, a macrocyclic but flexible system best seen as a "unimolecular reversed micelle." When **31a** dissolved in CCl_4 was stirred with water, ^{1}H NMR analysis revealed the presence of ca. 40 molecules of H_2O in the organic phase. However, when 3-M solutions of D-ribose **3**, D-glucose **9**, or D-fructose **23**

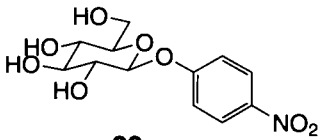
Table 4 Extractabilities of monosaccharide substrates from water into chloroform by receptor **24b**[a]

Substrate	Concentration of substrate in aqueous phase (M)		
	1.0	**0.5**	**0.1**
D-ribose **3**	0.7		
D-xylose **5**	1.1		
D-glucose **9**	1.0	0.5	<0.1[b]
D-galactose **10**	0.2	<0.1[b]	None detectable
D-mannose **11**	<0.1[b]	<0.1[b]	None detectable
Methyl-β-D-glucoside **14**	1.0		
Methyl-α-D-glucoside **15**	1.0		

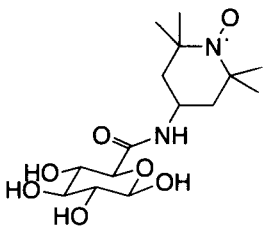
[a]Molar ratios sugar/**24b** appearing in the organic phase. $[\mathbf{24b}]_{org} = 2.9 \times 10^{-4}$ M.
[b]Carbohydrate detectable, but amounts too small for quantification by NMR integration.
Source: Ref. [32].



28

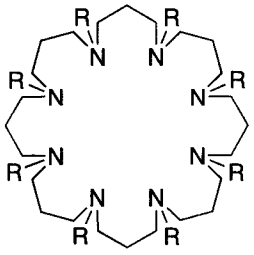


29



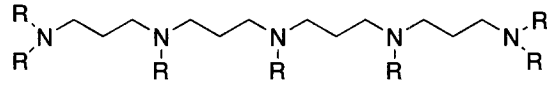
30

were substituted for water, the carbohydrate was extracted without any measurable coextraction of H_2O. In the case of D-glucose **9**, a 1:1 ratio of receptor/substrate was established directly by ^1H NMR integration, and also by reextraction of the sugar back into water. Sugar extraction, as well as water pool accommodation, were also observed with the acyclic heptakis(dihydroxyalkyl) reference host **32**, although no complexation was observed for **31b** or **33**, implying that the clustering of dihydroxyalkyl chains is essential for recognition.

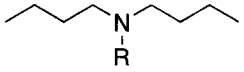


31a R = $CH_2CH(OH)CH(OH)(CH_2)_{10}CH_3$

31b R = $(CH_2)_{11}CH_3$

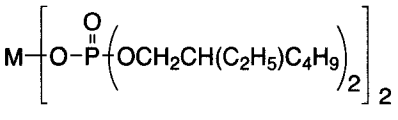


32 R = $CH_2CH(OH)CH(OH)(CH_2)_{10}CH_3$



33 R = $CH_2CH(OH)CH(OH)(CH_2)_{10}CH_3$

Kasuga et al.[35] have shown that the lipophilic alkaline earth metal salts **34** can promote transport of monosaccharides through chloroform bulk liquid membranes. Efficiency increased with cation size, the most effective carrier being **34** (M=Ba). A range of monosaccharides were tested as substrates, with the results shown in Table 5. Unsurprisingly, transport rates correlate to a fair degree with apparent substrate lipophilicity. However, the preference for D-ribose **3** over 2-deoxy-D-ribose **7** runs counter to this trend, and is especially interesting.



34 M = Mg, Ca, Sr, Ba

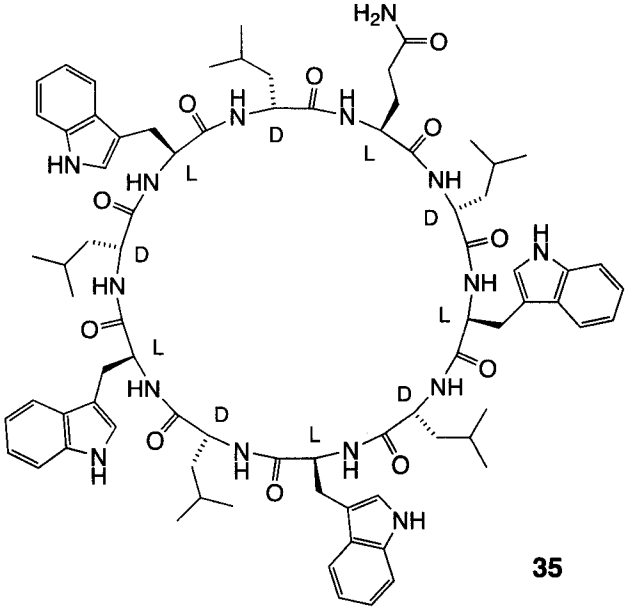
Table 5 Transport of monosaccharides by **34** (M=Ba)[a]

Sugar	Amount of sugar transported after 48 hr (mmol)
D-ribose **3**	3.0 (1.7)[b]
D-arabinose **4**	0.4
D-xylose **5**	0.3 (0.1)[b]
2-deoxy-D-ribose **7**	0.5
D-glucose **9**	0.1 (0)[b]
D-galactose **10**	0.1
D-mannose **11**	0.4 (0.2)[b]
L-fucose **12**	0.7
D-fructose **23**	0.7

[a][Receptor]$_{org} = 3 \times 10^{-2}$ M.
[b]Figures in parentheses refer to a competition experiment in which four monosaccharides were present simultaneously.
Source: Ref. [35].

CHANNEL-MEDIATED TRANSPORT OF CARBOHYDRATES

Finally, in the special case of aqueous phases separated by a lipid bilayer, polar substrates may be transferred via channels in the membrane. Bong et al.[36] have studied a range of self-assembling channels formed from cyclic peptides composed of alternating D and L amino acids. The peptides stack through hydrogen bonding to form a toroidal version of a β-sheet. If a decapeptide is used, the resulting tube has an internal diameter of ca. 10 Å, enough to allow passage of a glucose molecule.[37] The cyclodecapeptide **35**, composed mainly of lipophilic D-leucine



35

and L-tryptophan, was designed to concentrate in apolar environments. When added to liposome suspensions, it was shown by several spectroscopic techniques to incorporate into the bilayers to form H-bonded aggregates. If D-glucose **9** was entrapped in the liposomes, addition of **35** promoted efflux of the carbohydrate (measured using an enzyme-based assay). Kinetic studies were consistent with a channel, rather than carrier, transport mechanism. The analogous cyclooctapeptide, with an internal diameter of ~7.5 Å, showed no activity.

CONCLUSION

In summary, the study of carbohydrate recognition through noncovalent interactions has led to a range of phase transfer systems, effective for moderate to high concentrations of substrates. It remains difficult to transport monosaccharides into apolar media from dilute aqueous solutions. However, steady progress over the past 15 years suggests that the necessary affinities will be achieved before long. Success could have important consequences, not least in diabetes treatment where improved glucose sensors are sought with some urgency.[11–14]

ACKNOWLEDGMENTS

A.P.D. thanks all of his coworkers, whose names appear in the references, for their contributions to date and for their ongoing efforts in this area. Funding for our program on carbohydrate recognition has been provided by the EPSRC, the European Commission, the University of Bristol, and the Enterprise Ireland.

REFERENCES

1. Dwek, R.A.; Butters, T.D. Glycobiology—Understanding the language and meaning of carbohydrates. Chem. Rev. **2002**, *102* (2), 283–284.
2. Bertozzi, C.R.; Kiessling, L.L. Chemical glycobiology. Science **2001**, *291* (5512), 2357–2364.
3. Williams, S.J.; Davies, G.J. Protein–carbohydrate interactions: Learning lessons from nature. Trends Biotechnol. **2001**, *19* (9), 356–362.
4. Feizi, T.; Mulloy, B. Editorial overview: Progress at the frontiers of structural glycobiology. Curr. Opin. Struct. Biol. **2001**, *11* (5), 585–586.
5. Roseman, S. Reflections on glycobiology. J. Biol. Chem. **2001**, *276* (45), 41527–41542.
6. Weis, W.I.; Drickhamer, K. Structural basis of lectin–carbohydrate recognition. Annu. Rev. Biochem. **1996**, *65*, 441–473.
7. Sharon, N.; Lis, H. Lectins—Proteins with a sweet tooth: Functions in cell recognition. Essays Biochem. **1995**, *30*, 59–75.
8. Lee, Y.C.; Lee, R.T. Carbohydrate–protein interactions: Basis of glycobiology. Acc. Chem. Res. **1995**, *28*, 321–327.
9. Altamore, T.M.; Barrett, E.S.; Duggan, P.J.; Sherburn, M.S.; Szydzik, M.L. Cavitand boronic acids mediate highly selective fructose transport. Org. Lett. **2002**, *4* (20), 3489–3491.
10. Paugam, M.F.; Riggs, J.A.; Smith, B.D. High fructose syrup production using fructose-selective liquid membranes. Chem. Commun. **1996**; (22), 2539–2540.
11. Tohda, K.; Gratzl, M. A microscopic, continuous, optical monitor for interstitial electrolytes and glucose. ChemPhysChem **2003**, *4* (2), 155–160.
12. Katz, E.; Sheeney-Haj-Ichia, L.; Buckmann, A.F.; Willner, I. Dual biosensing by magneto-controlled bioelectrocatalysis. Angew. Chem., Int. Ed. Engl. **2002**, *41* (8), 1343–1346.
13. Wei, Y.; Dong, H.; Xu, J.G.; Feng, Q.W. Simultaneous immobilization of horseradish peroxidase and glucose oxidase in mesoporous sol–gel host materials. ChemPhysChem **2002**, *3* (9), 802.
14. Shafer-Peltier, K.E.; Haynes, C.L.; Glucksberg, M.R.; Van Duyne, R.P. Toward a glucose biosensor based on surface-enhanced Raman scattering. J. Am. Chem. Soc. **2003**, *125* (2), 588–593.
15. Davis, A.P.; Wareham, R.S. Carbohydrate recognition through noncovalent interactions: A challenge for biomimetic and supramolecular chemistry. Angew. Chem., Int. Ed. Engl. **1999**, *38* (20), 2978–2996.
16. James, T.D.; Sandanayake, K.R.A.S.; Shinkai, S. Saccharide sensing with molecular receptors based on boronic acid. Angew. Chem., Int. Ed. Engl. **1996**, *35* (17), 1911–1922.
17. Ayling, A.J.; Broderick, S.; Clare, J.P.; Davis, A.P.; Pérez-Payán, M.N.; Lahtinen, M.; Nissinen, M.J.; Rissanen, K. An extraction-based assay for neutral anionophores: The measurement of high binding constants to steroidal receptors in a nonpolar solvent. Chem. Eur. J. **2002**, *8* (9), 2197–2203.
18. Smith, B.D. Liquid membrane transport using boronic acid carriers. Supramol. Chem. **1996**, *7* (1), 55–60.
19. Cram, D.J. Preorganization—From solvents to spherands. Angew. Chem., Int. Ed. Engl. **1986**, *25* (12), 1039–1057.
20. *Comprehensive Supramolecular Chemistry*; Lehn, J.-M., Ed.; Pergamon: Oxford, 1996.

21. Diederich, F.N. *Cyclophanes*; Royal Society of Chemistry: Cambridge, 1991.
22. Vögtle, F. *Cyclophane Chemistry; Synthesis, Structures, Reactions*; Wiley: Chichester, 1993.
23. Aoyama, Y.; Tanaka, Y.; Toi, H.; Ogoshi, H. Polar host–guest interaction. Binding of nonionic polar compounds with a resorcinol–aldehyde cyclooligomer as a lipophilic polar host. J. Am. Chem. Soc. **1988**, *110*, 634–635.
24. Aoyama, Y.; Tanaka, Y.; Sugahara, S. Molecular recognition of sugars via hydrogen-bonding interaction with a synthetic polyhydroxy macrocycle. J. Am. Chem. Soc. **1989**, *111*, 5397–5404.
25. Kikuchi, Y.; Tanaka, Y.; Sutaro, S.; Kobayashi, K.; Toi, H.; Aoyama, Y. Highly cooperative binding of alkyl glucopyranosides to the resorcinol cyclic tetramer due to intracomplex guest–guest hydrogen bonding: Solvophobicity/solvophilicity control by an alkyl group of the geometry, stoichiometry, stereoselectivity, and cooperativity. J. Am. Chem. Soc. **1992**, *114*, 10302–10306.
26. Bonar-Law, R.P.; Davis, A.P.; Murray, B.A. Artificial receptors for carbohydrate derivatives. Angew. Chem., Int. Ed. Engl. **1990**, *29*, 1407.
27. Bhattarai, K.M.; Bonar-Law, R.P.; Davis, A.P.; Murray, B.A. Diastereo- and enantio-selective binding of octyl glucosides by a tetrahydroxycholaphane. J. Chem. Soc., Chem. Commun. **1992**, 752.
28. Bhattarai, K.M.; Davis, A.P.; Perry, J.J.; Walter, C.J. A new generation of ''cholaphanes'': Steroid derived macrocyclic hosts with enhanced solubility and controlled flexibility. J. Org. Chem. **1997**, *62*, 8463–8473.
29. Inouye, M.; Toshiyuki, M.; Masaru, F.; Hiroyuki, N. Molecular recognition of β-ribofuranosides by synthetic polypyridine–macrocyclic receptors. J. Am. Chem. Soc. **1995**, *117*, 12416–12425.
30. Inouye, M.; Takahashi, K.; Nakazumi, H. Remarkably strong, uncharged hydrogen-bonding interactions of polypyridine–macrocyclic receptors for deoxyribofuranosides. J. Am. Chem. Soc. **1999**, *121* (2), 341–345.
31. Davis, A.P.; Wareham, R.S. A tricyclic polyamide receptor for carbohydrates in organic media. Angew. Chem., Int. Ed. Engl. **1998**, *37* (16), 2270–2273.
32. Ryan, T.J.; Lecollinet, G.; Velasco, T.; Davis, A.P. Phase transfer of monosaccharides through noncovalent interactions: Selective extraction of glucose by a lipophilic cage receptor. Proc. Natl. Acad. Sci. **2002**, *99* (8), 4863–4866.
33. Greenspoon, N.; Wachtel, E. Reverse micelles as a model system for carbohydrate binding. J. Am. Chem. Soc. **1991**, *113*, 7233–7236.
34. Kobayashi, K.; Ikeuchi, F.; Inaba, S.; Aoyama, Y. Accommodation of polar guests in unimolecular polyamine–polyhydroxy cores: Solubilisation of sugars in apolar organic media via intramolecular polar microsolvation. J. Am. Chem. Soc. **1992**, *114*, 1105–1107.
35. Kasuga, K.; Hirose, T.; Aiba, S.; Takahashi, T.; Hiatani, K. Transport of monosaccharides through a liquid membrane mediated by lipophilic alkaline earth metal complexes. Tetrahedron Lett. **1998**, *39*, 9699–9702.
36. Bong, D.T.; Clark, T.D.; Granja, J.R.; Ghadiri, M.R. Self-assembling organic nanotubes. Angew. Chem., Int. Ed. Engl. **2001**, *40* (6), 988–1011.
37. Granja, J.R.; Ghadiri, M.R. Channel mediated transport of glucose across lipid bilayers. J. Am. Chem. Soc. **1994**, *116*, 10785–10786.

P

Photochemistry of Membrane-Coated Nanoparticles

Ulrich Siggel
Technischen Universität Berlin, Berlin, Germany

Guangtao Li
Jürgen-Hinrich Fuhrhop
Freie Universität Berlin, Berlin, Germany

INTRODUCTION

The objective of this article is to describe the different types of photochemical and photophysical reactions on nanoparticles and their application with special emphasis on recent work.

OVERVIEW

Nanoparticles can be used as single particles in solution (colloids) or arrays of particles with a specially designed architecture again in solution or on a suitable support. Membrane coating is often necessary for stabilization of aqueous or organic solutions of isolated particles. In particle arrays, special linker molecules may be used in order to realize an optimal interparticle distance.

Membrane-coated nanoparticles are functionalized for optical functions by covalent attachment of chromophores to the membrane or to the particle surface. Mixed monolayers of lipids and amphiphilic dyes are also common. The enclosed or adsorbed dyes are used for the construction of photochemical reaction centers as well as in the analysis of special membrane properties such as rigidity and regioselective interaction with solutes.

Photochemical reaction centers are realized either by covalent dyads or by noncovalent long-distance dimmers. Triads have also been realized. Nanoparticles provide a large surface to light-absorbing chromophores as well as the possibility to transfer electrons to electrodes. Functionalized membrane coatings allow the construction of long-distance dimers of redox active molecules for light-induced charge separation.

Three different types of colloidal particles are currently used: metals, semiconductors, and insulators. Colloids of noble metals (gold, silver, copper) are characterized by their chemical stability and by surface plasmon absorptions in the visible range, which interfere with the photophysics of dyes. Semiconductor particles with a narrow bandgap absorb in the visible region and are prone to corrosion. Semiconductors with a wide bandgap, on the other hand, are more stable, but absorb only in the UV region. Dyes absorbing visible light are then applied as sensitizers in order to use sunlight. Insulating particles are not appropriate for the construction of photochemical cells but are promising carriers for studies of charge separating molecular assemblies.

GOLD AND OTHER METAL PARTICLES

Gold particles have been used most frequently. They can be easily prepared by different methods. The citrate reduction method[1] yields nearly monodisperse spheres with typical diameters of about 20 nm. Particles of this size remain dispersed for many days and their curvature is low enough to allow the construction of closed membrane systems in water. Weakly bound citrate leads to a negative surface charge. Particles of silver, palladium, and platinum may be prepared by the same method.[2–5] Borohydride reduction, on the other hand, yields much smaller particles (1–5 nm). Their high curvature does not allow the formation of a closed membrane. Monolayer-protected gold clusters (MPCs) are obtained by reaction with alkane thiols, which form a covalent bond to the gold surface. The surface of the nanoparticles can be varied from hydrophilic to hydrophobic by selecting appropriate head groups on the other end of the attached thiol.[6,7]

Other routes of particle production include the electrochemical synthesis of size-selective nanostructured palladium clusters in tetrahydrofuran[8] and a photochemical method of making size-controlled spherical metal particles in a nonaqueous, but highly polar and viscous medium. Complexes of gold, silver, palladium, and platinum have been reduced by the photochemically formed carbamoyl radicals.[9] Small particles were sterically stabilized by polyvinylpyrrolidone (PVP), the size being determined by the ratio PVP/metal complex. A ratio of 5/1 yielded, for example, particles of 12 nm (Au), 15 nm (Ag), 10 nm (Pt), and 16 nm (Pd). The gold particles were pure enough to show single-electron tunneling (SET) in the current–voltage dependence (the 1.8-nm stabilizer shell

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009101

serves as a tunnel junction). This makes the particles interesting for nanoscale electronic circuits.

The functionalization of the monolayer opens the way to a variety of photophysical and photochemical reactions, which are strongly influenced by the collective excitation of metal electrons, leading to the size-dependent plasmon absorption. This leads, however, to heating of the probe upon strong excitation, problems in the measurements of transient absorptions of adsorbed dyes, energy transfer to the metal, and reduced quantum yields for all the reactions of the chromophore from the excited singlet state. Analytically, the plasmon absorption has been used to probe the dielectric properties of surface-bound molecules,[10] and the distance between coupled particles if aggregation of the particles occurs.[11]

Photoreactions have been useful in exploring the properties of monolayers on planar gold surfaces. Photoextrusion of nitrogen from aryl azide[12,13] and diazoketones,[14,15] E-Z photoisomerization of stilbenes[16] and azocompounds[17] as well as photodimerization reactions[18–21] have been well characterized. On gold nanoclusters electron transfer reactions are very often examined; they occur from donors either to relay molecules or directly to the metal. *cis–trans* Isomerization of molecules with double bonds and cycloadditions have, however, also been studied. The latter are very much enhanced if the reporter molecules are close to each other and in proper orientation. If dimerization, followed by charge transfer, occurs in the excited state, excimer fluorescence may be studied as a measure of distance. As a further photophysical reaction, energy transfer can also be used to probe the membrane order.

Energy Transfer

Singlet energy transfer is measured by fluorescence quenching of suitable chromophores. The plasmon absorption band of metallic particles is usually broad and energy transfer occurs with a wide range of fluorescent dyes. The transfer from donors to acceptors depends on the distance: the efficiency of a transfer between a point dipole and a nearby flat metal surface decreases with the third power; for extended molecules close to curved particle surfaces with an intermediate exponent between 3 and 6 is expected. Monolayers on small particles are usually fluid and the chain density decreases with increasing chain length because of the high curvature. Motional freedom, thermal flexing, and chain folding lead to altered energy transfer efficiencies.

Imahori et al.[22] have shown that the energy transfer between tetraphenylporphyrin, coupled to a gold surface by a C12-alkanethiol spacer, and gold is much more effective in the case of a SAM on planar gold (111-surface) as compared to particles of 5-nm diameter (MPC). The lifetime of the porphyrin fluorescence is shortened from 9.5 nsec to 0.04 and 0.15 nsec, respectively. Additionally a 9.1-nsec phase (15% of total amplitude) appears on the particles, suggesting a second type of monolayer structure. The portion of unquenched porphyrin molecules is only 3% for mixed monolayers. Lifetime measurements thus result in additional information on the homogeneity of the monolayers. Energy transfer is more efficient in SAMs than in MPCs. The origin of this effect is not known.

Aguila and Murray[23] have investigated the energy transfer of dansyl chromophores bound to gold clusters of 5–8 nm (C2 to C15 linkers) in mixed monolayers. In this environment the fluorescence is reduced to 1% of that of the free dye by energy transfer. The emission intensity increased with the effective distance, provided that the linker had about the same length as the surrounding alkanethiolates (C4–C12). For longer linkers the emission decreased and became dependent on the load with the dansyl label, i.e., the stiffness of the outermost part of the disordered layer (Fig. 1).

The distance-dependent energy transfer from a methylene blue-type fluorophore (*Alexa 488*) to a metal surface[24] has been exploited to measure changes of chemisorption with respect to the electrode potential on a polycrystalline gold electrode. In the potential range of -0.7 to $+0.2$ V only reversible reorientation of the charged chromophore occurred. For potentials lower than -0.8 V and higher than $+0.2$ V, an exponential increase of fluorescence was observed. This was related to the reduction and oxidation of the Au–S bond and release of the fluorophore into the solution. An additional spacer of 12 C-atoms between the chromophore and the thiolate moiety decreased the sensitivity toward the potential, because the van der Waals forces between the alkane chains became dominating.

Excimer Fluorescence

The excimer fluorescence of pyrene is a good tool to investigate order and dynamics of alkane monolayers on surfaces. Excimers are not formed in homogeneous solution of low concentration. The fluorescence behavior after the addition of a gold colloid depends on the nature of the pyrene side chain. If this does not contain a thiol group for binding to the gold the monomer fluorescence is somewhat decreased because of ineffective energy transfer to the gold. On binding through a thiol group hydrophobic aggregates are formed and a strong excimer fluorescence appears. This has been shown by Chen and Katz[25] for protected thiols. A pyrene-thioester and a pyrene-thiocarbonate bind noncovalently to the surface of 12.5-nm citrate-gold particles with binding constants of 7.2×10^7 and 1.7×10^8 M^{-1}, respectively. The change from excimer to monomer fluorescence has also been

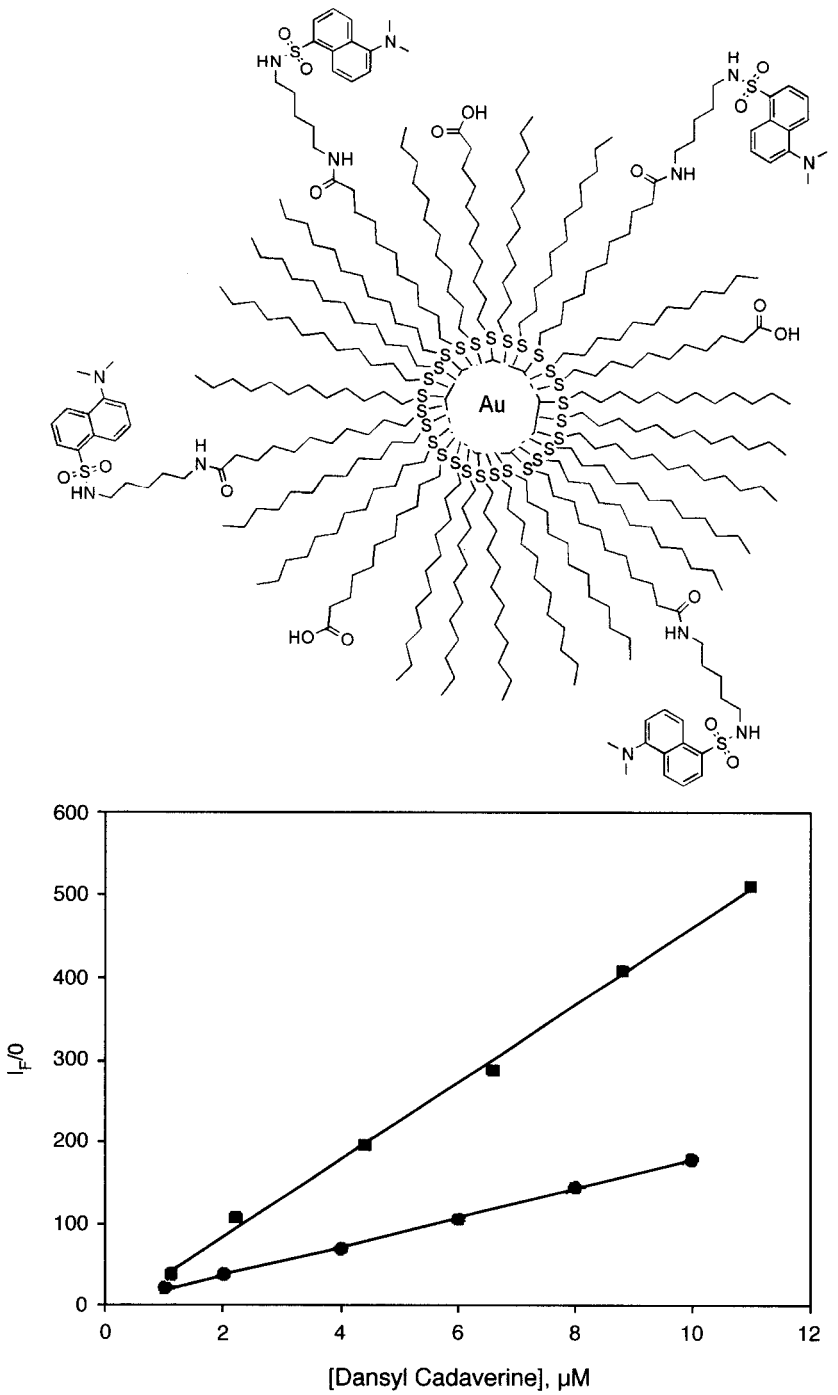


Fig. 1 Monolayer protected cluster ("MPC") of an HSC_{10}-coated Brust gold particle after partial exchange with $HSC_{10}COO^-$ and coupling with dansyl cadaverine and fluorescence intensities with 17 (upper curve) and 6 dansyls (lower curve) per MPC. Fluorescence increases linearly with concentration in both cases. The dye molecules do not interact. (From Ref. [23].)

taken as a measure for the gold-catalyzed hydrolysis of the ester and carbonate bonds, which leads to the appearance of the pyrene moieties in the bulk solution.

Pyrene excimer fluorescence has also been used to detect changes in interparticle interactions for capped gold particles.[26] Freshly prepared Brust-type gold particles (3.2–3.7 nm) capped with pyrenyl-decanethiol or -heptadecanethiol showed practically no excimer fluorescence. The monomer fluorescence was quenched by 99% by energy transfer to the gold particles. Although 43% (C-10)

and 31% (C-17) of the surface gold atoms were covered with the thiol, most of the pyrene units were not oriented in a face-to-face manner. Upon ageing for up to 300 hr the fluorescence increased, especially the excimer emission. This was traced back to aggregation of the gold particles, leading to interdigitation of the alkyl chains resulting in a stacking of the pyrene units. The higher intensity of the excimer fluorescence for the short-chain pyrene provides further evidence for the chain density gradient for alkanethiolates on gold nanoparticles. The decay of the fluorescence was biexponential with lifetimes of 4.3 and 22 nsec for the monomer fluorescence of fresh solutions. For aged samples only the lifetime of the longer-lived excited species was increased to 35 and 30 nsec for the two preparations. This indicates that there may be different binding sites on the gold surface leading to different excited states (Fig. 2).

trans–cis Isomerization, Cyclodimerization, and Benzylic Cleavage

The realization of dimerization necessitates two molecules in near neighborhood; isomerization, on the other hand, needs some space around the molecule. Both reactions are optimized under opposite conditions. Fluid monolayer coatings favor cleavage reactions which are linked to release molecules.

The comparison of three photochemical reactions, *trans–cis* isomerization, cyclodimerization, and benzylic cleavage of thioalkyl-azobenzenes and -stilbenes in

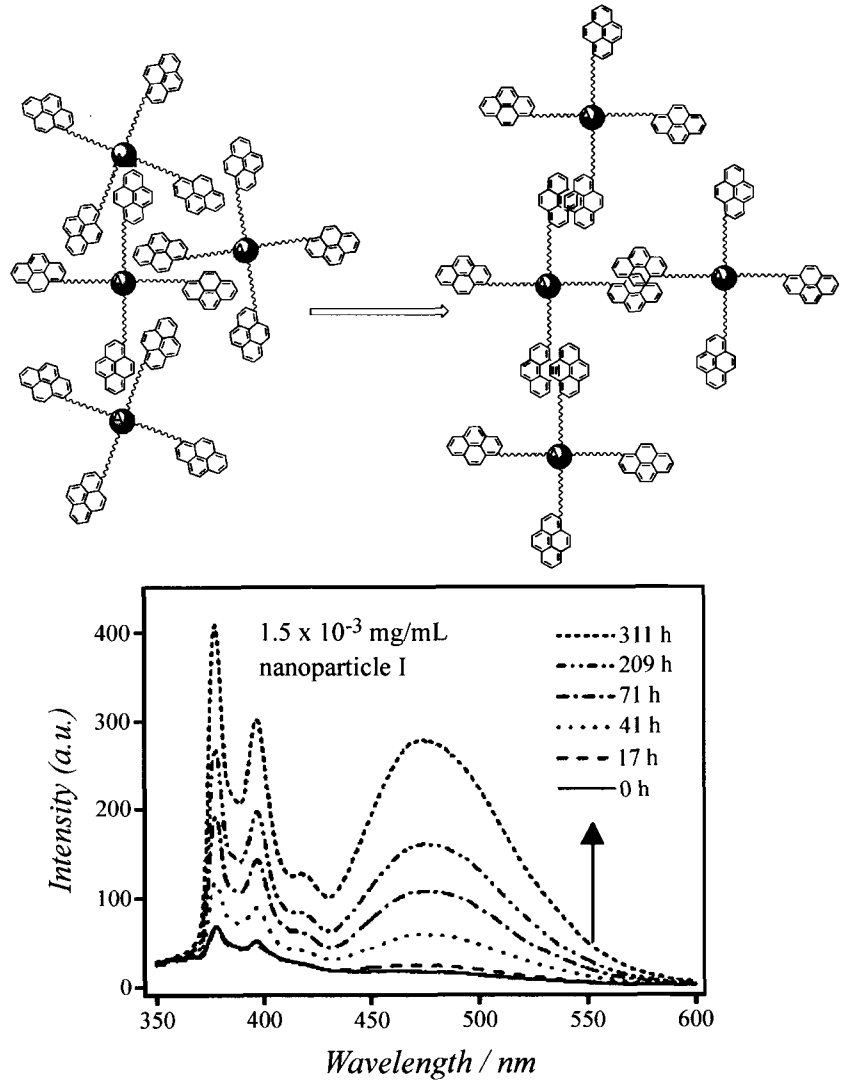


Fig. 2 Fluorescence of pyrene-coated particles increases with ageing because interdigitation leads to excimers. (From Ref. [26].)

homogeneous solution and in monolayers on 2.5-nm gold particles and on planar gold,[27,28] yielded valuable information about the effects of monolayer packing. Photoisomerization was found to be independent of the alkyl chain length in homogeneous solution with quantum yields of 10^{-1} and 3×10^{-4} for the azobenzene and stilbene species, respectively. In a tightly packed monolayer on planar gold it was sterically blocked. In the more loosely packed monolayers on the small gold particles, isomerization occurred with reduced quantum yield. An additional decrease of the yield was found after shortening the alkyl chain from 12 to 4 CH_2 groups. Quenching of the excited state by the metallic core became dominant. The dependence on the distance was exponential with an attenuation constant of β=0.5, which is smaller than expected for through-bond coupling. If a nitrobenzoxy group was introduced into the alkyl chain, photocleavage occurred. Parallel observation of isomerization and cleavage pointed to strong metal-based nonradiative quenching of the excited state. The significance of packing and steric hindrance became evident in stilbene photodimerization experiments. It was not observed in solution or in monolayers on planar gold but occurred on gold particles with a very low quantum yield of $2\text{–}4\times10^{-6}$, depending on the distance (Fig. 3).

Similar conclusions with respect to the rigidity of the monolayers have been drawn by Evans et al.[29] Photoswitching between the *cis* and *trans* states of alkane-derivatized azobenzene was not possible in single-component monolayers on planar gold. It has only been realized in mixed monolayers with free volume introduced by a spacer molecule with three ethyleneoxy units and especially in mixed monolayers of C6-thiol-azobenzene and mercaptophenol on 3.2-nm particles.

The Norrish–Young type II photochemical reaction has been used to probe the properties of a variety of ordered media including cyclodextrins, zeolites, liquid crystals, and micelles. At first the reaction was investigated in an MPC capped with a mixed monolayer of short alkanethiol and mercaptoundecanophenone by Kell et al.[30] Irradiation released free benzophenone, irrespective of the length of the spacer alkanethiol (C-6 or C-12), but not in its absence. Whether fragmentation or cyclization is the main pathway to products depends greatly on the conformational flexibility of the triplet 1,4-biradical, which is formed from the n,p* triplet excited state of the carbonyl

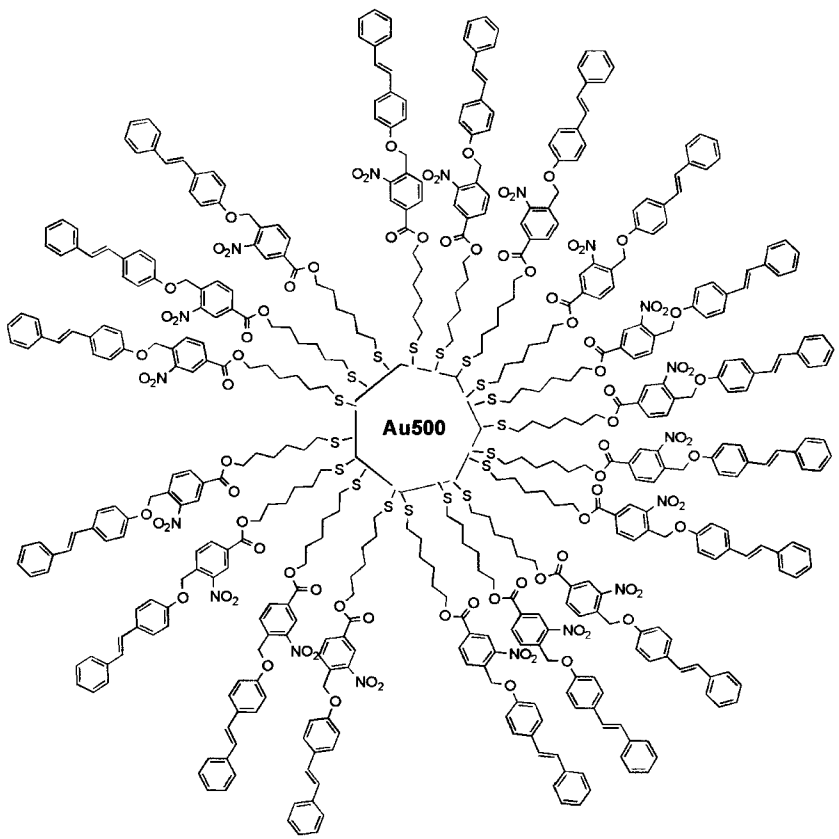


Fig. 3 Photodimerization of stilbene chromophore was only observed in ordered and flexible domains on particles with high curvature. (From Ref. [28].)

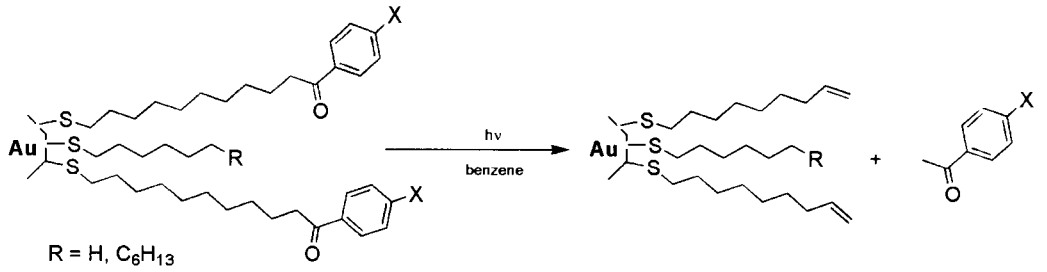


Fig. 4 Benzophenone is only released, if the benzophenone head groups are separated by unreactive molecules and on curved surfaces. (From Ref. [30].)

group. If the volume, which is necessary for cyclization motions, is not available, only fragmentation occurs. The resulting terminal vinyl group on the MPC can then be used for further modification (Fig. 4).

Miscellaneous Reactions

An interesting example for the interaction of a chromophore with the metal particle is the binding of methylaminopyrene on spherical gold particles of 5–8-nm diameter.[31] Because of internal charge transfer, pyrene is only weakly fluorescent in THF solution. On binding to the gold surface the electron-donating ability of nitrogen was decreased. As a result the fluorescence yield was increased from 0.09 to 0.48 (chelation enhancement). A second indication of the binding was the appreciable red shift of the absorption and fluorescence bands.

A phthalocyanine thiol bound to Brust-type gold particles with diameters ranging from 2 to 4 nm and associated with the TOAB phase transfer reagent generated singlet oxygen upon irradiation with visible light more effectively than without gold. The quantum yield increased from 0.45 to 0.65. TOAB promotes solubility in polar solvents, which is essential for a possible application in photodynamic therapy.[32]

Upon illumination of thiocyanine (TC)–J aggregates in solution, an ultrashort excited state was formed. Its relaxation was intensity dependent and dominated by exciton–exciton annihilation. On bulk silver surfaces ultrafast quenching of the exciton and its fluorescence occurred. The same was true on gold nanoparticles. The binding of the dye aggregate via two sulfonate groups was reported originally by Kometani et al.[33] who applied gold, silver, and composite Au/Ag particles with diameters of about 10 nm. The absorption spectra were measured and simulated using the Maxwell–Garnett treatment. On silver particles after fsec-excitation at 417 nm (140-fsec time resolution) a charge separated state with a lifetime of 300 psec was observed[34] by absorption changes at 475 nm close to the sharp absorption band of the J aggregate on silver at 481 nm. The exciton of the J aggregate is then of course strongly coupled to the silver particle plasmon. Photoexcitation of this resonance leads to very fast electron transfer of the TC aggregate to the metal. In the case of gold particles the plasmon energy is too low for such coupling (2.38 vs. 3.1 eV for Ag).

Rhodamine 6G has been used as a reporter molecule for changes in aggregation and morphological changes of the gold clusters. The 2-nm particles, prepared by rhodanide reduction,[35] did not exhibit surface plasmon absorption, for which the diameter must be at least 5 nm,[36] and bound rhodamins 6G as a monomer. On increasing the Au/dye ratio, aggregation was shown by a blue shift of the dye absorption (H aggregate) and the appearance of a plasmon band at 537 nm. Twenty-five minutes of laser irradiation at 532 nm lead to melting and growth of the particles, due to plasmon excitation.[37] This was indicated by the broadening of the plasmon band and by an increase of the monomer fluorescence of rhodamine at 545 nm, originating from molecules released into the bulk solution (Fig. 5).

Well-Defined Assemblies of Nanoparticles

Arrays of nanoparticles have been prepared by salt-induced aggregation,[38–42] Langmuir techniques,[43] surface assembly,[44–46] or crystallization.[47] The motive for such preparations is the special nonlinear optical properties of collective metal nanoclusters. They show large surface-enhanced Raman signals[44–46] and large first hyperpolarizabilities.[38–42] A necessary condition is a well-defined symmetry and interparticle spacing, which is not always met in the preparations mentioned above. It may be realized more easily by covalent bridges between the particles.[48,49] Gold particles were connected, for example, by thiol-functionalized oligo-phenylacetylenes.[50] The intensity of hyper-Rayleigh scattering (HRS), i.e., incoherently scattered second harmonic light, was especially high for noncentrosymmetric trimers, which was traced back to enhanced values of first hyperpolarizability. These properties are also of interest for nanoscale electronics.

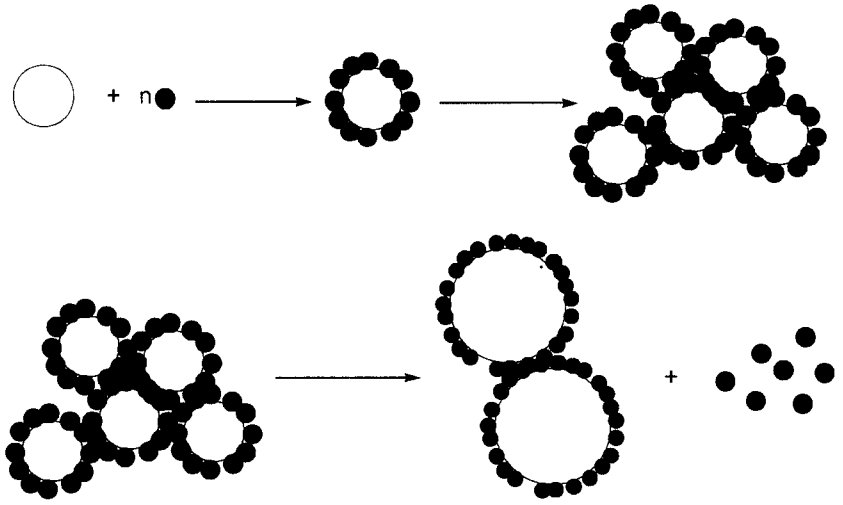


Fig. 5 Model of possible morphological changes of gold-dye cluster arrays upon laser irradiation. (From Ref. [37].)

Defined aggregates can also be obtained by noncovalent interactions. Two-dimensional (2-D) assemblies with hexagonal packing of 5-nm gold particles (Brust-type) were formed by adsorption of tetrapyridylporphyrin to the surfaces. Optically, this was shown by the red shift (512 to 610 nm) and broadening of the plasmon absorption as well as of the porphyrin Soret band from 25 nm half width in solution to 45 nm. The porphyrin dye acts as a reporter molecule and as a trigger of a defined aggregation.[51] Surface-enhanced resonance Raman scattering (SERRS) was also observed (Fig. 6).

Photoelectrochemical Cells

A photoelectrochemical cell consists of a charge separating system attached to a macroscopic electrode and a redox electrolyte carrying charges to a counter electrode. If self-assembled monolayers (SAMs) of electron transfer dyades or triades are directly bound to the working electrode, the light absorption in these 2-D arrays is normally low. The use of photosensitive nanoparticles in contact with the electrode yields 3-D array electrodes with high surface and high absorbance.

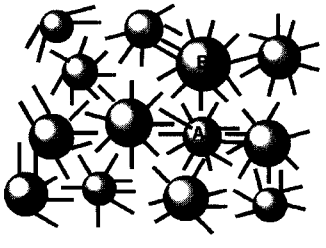


Fig. 6 H_2TpyP stands perpendicular on the gold surface and causes aggregation. (From Ref. [51].)

The conventional method to get a 3-D array of a charge separating system makes use of bifunctional bridge compounds and of electrostatic interactions to form a multilayered system. Lahav et al.[52,53] used indium-doped tin oxide (ITO) electrodes coated by aminosiloxane. Negatively charged gold nanoparticles (diameter: 13 nm) were self-assembled on the ITO electrode, followed by a positively charged charge-separating (CS) system. Repetition produced a multilayer system. The CS system consisted of either a Zn protoporphyrin–viologen dyade or a catenane, consisting of a cyclo-bis-viologen-phenylen and a Ru(bipy) derivative. Following excitation of the porphyrin the excited electron is conducted to the ITO electrode via the gold particles. The quantum yields of the corresponding photochemical cells [with ethylenediaminetetraacetic acid (EDTA) as electron donor and a graphite counter electrode] were $\phi = 10^{-3}$ ($\lambda = 544$ nm) for the porphyrin system and $\phi = 10^{-4}$ ($\lambda = 434$ nm) for the Ruthenium system. These numbers, referring to the short-circuit current, are relatively low. This is because of the poor contact between the CS system and the gold particles (Fig. 7).

Better results are obtained if the photosensitizer is directly bound to the gold particles. Sudeep et al.[54] electrodeposited functionalized gold particles on a nanostructured SnO_2 electrode, made by annealing a SnO_2 film on an optically transparent electrode (OTE). The charge separating system consisted of a light-absorbing fullerene electron donor bound to 3-nm gold particles via a modified alkanethiol chain. The light-induced electron transfer to the gold particles was optimized by the redox electrolyte J^-/J_3^- in high concentration. An incident

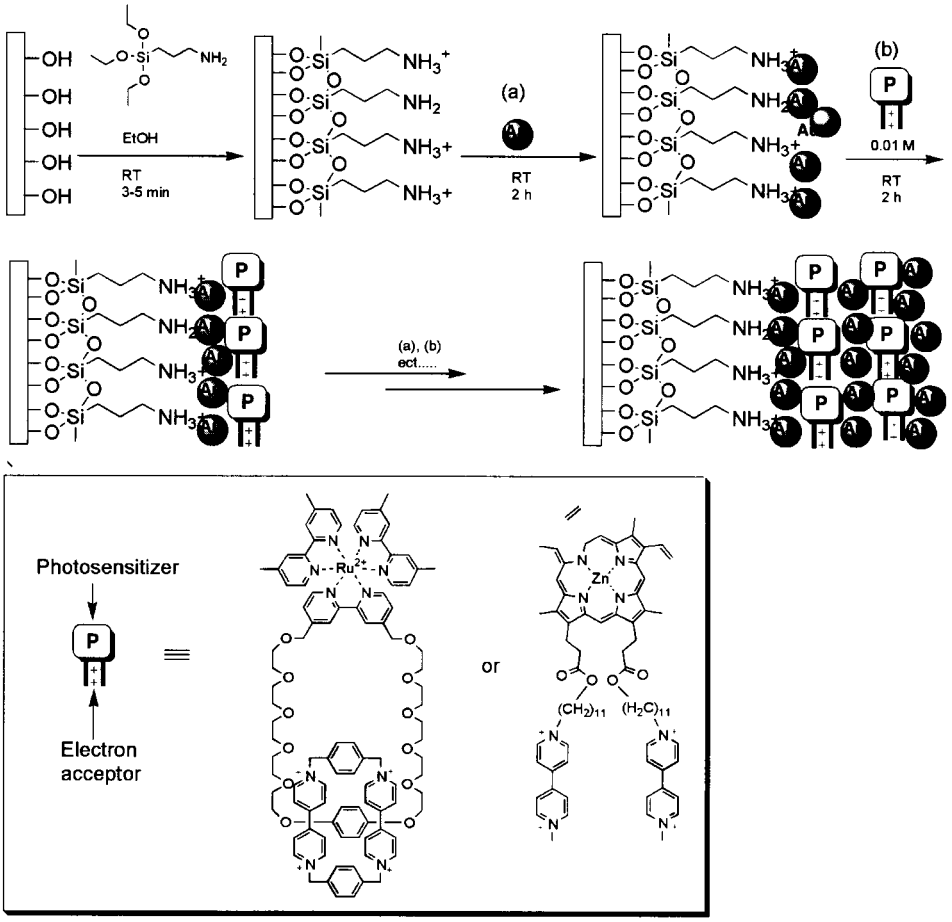


Fig. 7 Model of the gold-nanoparticle arrays on an ITO electrode for photocurrent production. (From Ref. [53].)

photon to photocurrent efficiency (IPCE) of 8×10^{-3} (at 450 nm) was achieved (Fig. 8).

Another efficient approach is to precipitate the gold particles directly on a flat gold electrode and to bind an electron transfer dyade afterwards.[55] The dyade consisted, for example, of a rutheniumbipyridyl complex as light-absorbing electron donor, connected via a C7 bridge to viologen, as electron acceptor, and a second bridge (C6) to the final thiol group bound to the citrate gold particles. In such a photochemical cell with triethanolamine as electron donor and a Pt-counter electrode the photocurrent was 15 times higher than in a cell in which the RuVS dyade was directly bound to the flat gold electrode. This factor corresponds to the enhanced concentration of RuVS as achieved by binding to the nanoparticles. The photocurrent efficiency was 8×10^{-3} (at 460 nm).

High IPCE was, however, also achieved on flat gold electrodes without nanoparticles by increasing the absorption cross section of the sensitizing electron donor. Imahori et al.[56] prepared mixed SAMs of a ferrocene–porphyrin–fullerene electron transfer triad and of a boron–dipyrrin sensitizer, which transfers energy to the porphyrin. An IPCE as high as 1.6×10^{-2} (at 430 nm) was thus achieved.

Construction of Form-Stable Gaps in Monolayers

Li and Fuhrhop[57] used a totally different approach for capping and functionalizing gold particles. Multiply charged derivatives of tetraphenylporphyrins were directly bound to the gold surface in a flat-lying position. Long-chain thiols with two secondary amide groups were then self-assembled around them. As a result of hydrogen bonding the monolayers were rigid and impermeable to compounds, which dissolve in fluid membranes. The diameter of above 20 nm of the particles allowed the formation of totally closed membranes. The curvature was not too high. Form-stable gaps of $2 \times 2 \times 2$-nm size with rigid walls were thus formed. The rigidity has been proven by fluorescence quenching experiments. The porphyrins on the surface showed only a very weak rest fluorescence

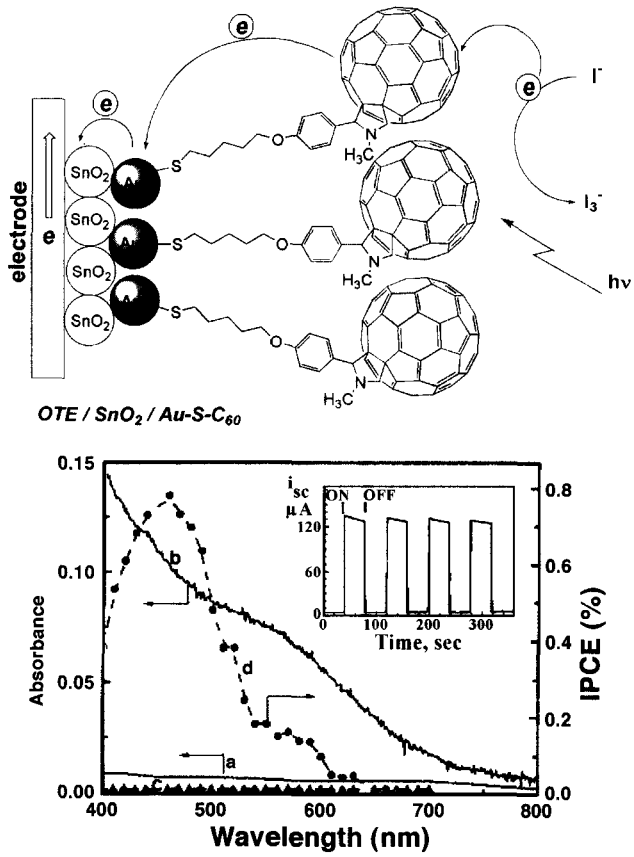


Fig. 8 Model of fullerenethiol-coated gold nanoparticles, their absorption spectrum (–) and photocurrent action spectrum (- - - -). The inset shows the on–off response upon illumination at 400 nm. (From Ref. [54].)

as a result of energy transfer quenching. It was totally quenched by paramagnetic metalloporphyrins of the same size as the porphyrin at the bottom of the gap by forming a heterodimer. Porphyrins with larger side groups were not able to enter the rigid gaps and quench the fluorescence. Such size discrimination did not take place when the surface layer around the porphyrins consisted of octadecanethiol without rigidifying amide groups. On flat gold electrodes, long-distance (1–2 nm) heterodimers were also established with a fitting porphyrin being attached to amine substituents at the rigid walls of the gap[58,59] (Fig. 9).

SEMICONDUCTOR NANOPARTICLES

Semiconductors can be optically excited. But the internal charge separation is not stable. The electrons in the conduction band and the holes in the valence band have to be brought to reaction with added electron acceptors and donors, respectively. In this way charge separation can be stabilized. Much of the work with semiconductor particles is thus focused on the conversion of sunlight into chemical energy (e.g., hydrogen production) or electrical energy (photoelectrochemical cells). For the absorption of light the size of the bandgap is essential. The coupling to

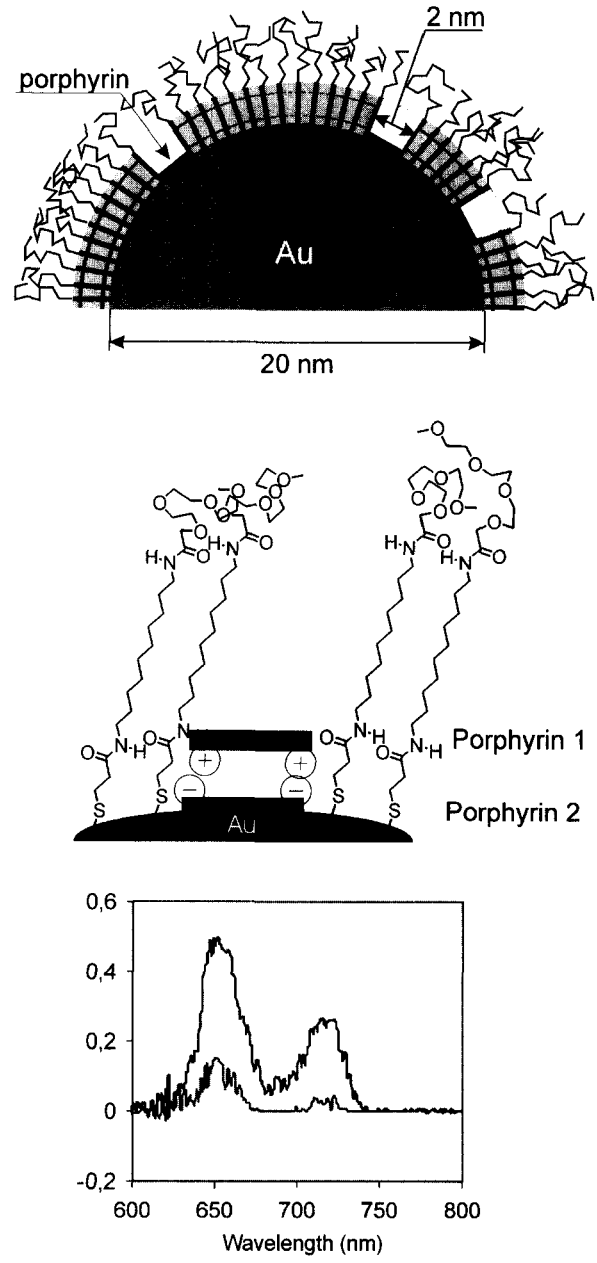


Fig. 9 Model of a gold nanoparticle coated with a rigid diamide monolayer containing 2-nm gaps with a porphyrin at the bottom. The head group of the diamido amphiphile is an oligoethylene ether. (From Ref. [57].)

electron acceptors and donors depends on the position of the conduction and the valence band. It must be noticed that these properties are size dependent in the quantum (Q) particles.[60] Narrow bandgap semiconductors such as cadmium sulfide (CdS),[61] cadmium selenide (CdSe),[62] lead silfide (PbS),[63] and silver sulfide (Ag_2S)[64] are excitable with visible light. Wide bandgap semiconductors such as titanium dioxide (TiO_2),[65] zinc oxide (ZnO),[66] and tin dioxide (SnO_2)[67] are only sensitive in the UV. Spectral sensitization is therefore an important approach to making wide band semiconductors attractive for solar energy conversion.[68,69] A chromophore absorbing in the visible range is attached to the particle surface. The excited electron of the chromophore is injected into the conduction band of the semiconductor. Thus in most cases the semiconductor surface is solely modified by dyes; additional monolayer coatings may improve solubility and longevity in the chosen solvent.

A different approach to enlarging the wavelength range of absorption is the coupling of two semiconductor particles with wide and narrow bandgap in mixed particle assemblies.[70] Either nanostructured films containing nanoparticles of different semiconductors (e.g., ZnO and CdS) are prepared by coprecipitation.[71] Heterostructures of the core-shell type are formed by controlled precipitation of semiconductor molecules of one type (shell) on the presynthesized nanoparticles of another type (core), e.g., CdSe/ZnS.[72] Photocatalytic water splitting with such systems was introduced by Henglein[73] and Duonghong et al.[74] both in 1982. It should be noted that in most cases the semiconductor particles used in photocatalytic studies (synthesis of organic compounds) are rather large (100 nm) with no quantum size effects.

Spectral Sensitization

Charge ejection from excited dyes or charge injection into the conduction band of semiconductor particles is the key reaction for the construction of photoelectrochemical cells of technical interest. At first, electron injection was not determined directly. Photocurrents were measured in electrode systems with polycrystalline layers of TiO_2 loaded with different dyes.[75–80] From the current efficiency IPCE and the absorption of the dye quantum yields of injection from 87% to 100% were calculated. Direct measurement of the luminescence yield of the ruthenium complex on TiO_2 particles gave a value of only 60%.[75] Later,[81] fluorescence and transient absorption changes were measured on colloidal TiO_2 electrodes, consisting of SnO_2 conductive glass and a film of 25 nm TiO_2 particles. Chlorin e6 and Cu-chlorophyllin were used as sensitizers. Comparison of the fluorescence lifetime of chlorin e6 on TiO_2 and ZrO_2, where charge injection is not possible, yielded a rate constant for the electron injection of 2.2×10^9 sec^{-1}. The constant was nearly 10 times smaller (3×10^8 sec^{-1}) for the copper compound, where electron transfer starts from the tripdoublet state and was determined from the phosphorescence lifetimes. Transient absorption changes after 10 nsec of excitation of the dye indicated the formation of its cation radical and of the electron injected into the TiO_2 conduction band. In the absence of a redox electrolyte (J^-) the recovery of the ground-state absorption of the dye and the decay of the products of charge separation were identical. The addition of J^- caused a fast reduction of the cation radical, whereas the absorption of the electron stayed for more than 100 msec.

Fsec-time resolution is necessary to the study of the dynamics of electron injection. This was done for fluorescein 27 adsorbed to 2.4-nm TiO_2 particles.[82] The absorption of the injected electrons had a rise time of 300 fsec, which is identical to the decay time of the stimulated emission of the dye. The recombination was much slower, which was explained by the validity of the Marcus formula in the inverted region. The kinetics was multiexponential with a very wide distribution of rates, ranging from 25 psec to nanoseconds and microseconds. Normally, in such a case, a distribution of energetically different trap sites for electrons is suggested. Here a red shift of the transient absorption within 25 psec was found, explained by vibrational cooling and perhaps solvation and structural relaxation of the initial product state, leading to a time-dependent ΔG^0.

Ashbury et al.[83] studied the effect of the bridge length of the binding ligand of re-polypyridyl complexes. They also obtained evidence for the injection of hot excited state electrons.[84] Bonhote et al.[85] studied the structural effect of different Ru-dye molecules. The best yield of charge separation was obtained for the system in which the excited electron was localized on the ligand bound to the semiconductor surface.

Electron injection is expected to be especially effective from a highly polar excited state of the dye. This has been verified by comparison of two coumarin dyes bound to TiO_2 particles through a carboxyl group.[86] Coumarin C 343 is a four-ring-system and the electron injection occurs from a normal intramolecular charge transfer (ITC) state with a quantum yield of 0.6. In the second dye, 7-diethylaminocoumarin-3-carboxylic acid (D-1421), the ICT state was able to relax to the more polar twisted intramolecular charge transfer (TICT) state. The overall quantum yield of electron injection from both excited states was 0.9, i.e., 50% higher than for C343, for which rotation of the nitrogen moiety is not possible. The products of the reaction, the dye radical cation, and the electron in the TiO_2 conduction band were shown by picosecond laser flash photolysis.

Photocurrents in Photoelectrochemical Cells

For a long time the construction of efficient photoelectrochemical cells was hampered by the fact that monolayers of dyes on planar electrodes absorb usually less than 1% of the incident light. Graetzel[75–80] used polycrystalline TiO_2 (anastase) electrodes with a roughness of 200. Thus in the beginning the second problem, that of charge injection, was attacked. $Fe(CN)_6^{4-}$ which forms a charge transfer complex with TiO_2 yielded an IPCE of 0.37 only;[76] the metalloporphyrin ZnTPPC[77] was only slightly better (0.42). A number of ruthenium complexes were examined.[87] The normal *tris*-bipyridin complex did not bind sufficiently. Carboxylic groups turned out to be necessary. Ruthenium complexes with two and three carboxy-bipyridin ligands were used and gave values of 0.62 and 0.73 for the IPCE.[75,78,79] For the latter ligand the importance of type and concentration of the redox electrolyte were determined.[79] J^- turned out to be better (0.73) than Br^- (0.56); and for hydroquinone the initially used concentration of 1 mM was too low (0.44) and had to be increased to 0.1 M (0.74). A very good value (0.83) was obtained with coumarin.[80] Other groups also tried merocyanines.[88] The importance of carboxyl groups for binding of a dye to the TiO_2 surface was again shown by comparing the photocurrent action spectra of chlorophyll, chlorins, and mesoporphyrins.[89] Copper-chlorophyllin was not better than copper-chlorin, indicating that conjugation of the carboxyl group to the π-electron system of the dye was not advantageous.

O'Regan and Graetzel published in 1991 the design of the first photoelectrochemical cell which could compete with photovoltaic cells.[90] The progress was based on the use of a 3-D network of TiO_2 nanoparticles (15 nm size), coated with a charge-transfer dye and fixed as a film of 10-μm thickness on a conducting glass support. The large surface allowed binding of a higher amount of the dye than before. The dye was a trimeric ruthenium complex RuL_2 (μ-(CN)Ru(CN)L_2')$_2$, with bipyridine and bipyridine-dicarboxylic acid as ligands. With lithium-iodide as a redox electrolyte an IPCE of 0.84 was achieved. The overall energy conversion yield was 7.1–7.9% in simulated solar light and 12% in diffuse daylight. The next step was to improve the redox electrolyte by changing to solid systems. P-type semiconductors[91] and organic materials[92] were tested with disappointing results. An amorphous organic hole transporting material, namely, tetrakis (*N*,*N*'-di-methoxyphenyl-amine)-spirobifluorene (OMeTAD), was then successful.[93] The energy efficiency was 33% with Ru-bis(dicarboxy-bipyridyl)-bis-rhodanid complex as a sensitizer (Fig. 10).

CdS nanoparticles were covered with thiolated oligonucleotides and then cross-linked with DNA. The assembly was connected with a gold electrode by double-

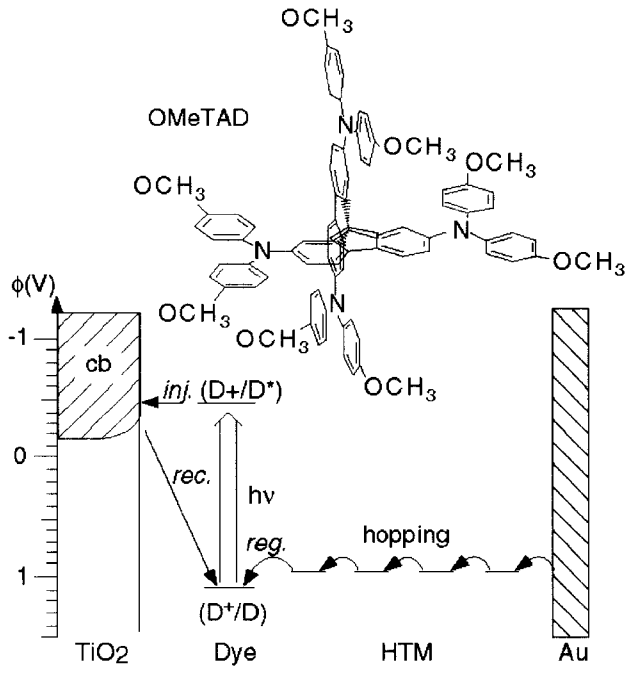


Fig. 10 Model of the electron injection (inj.), regeneration (reg.), recapture (rec.), and hopping in the mesoporous TiO_2 solar cell heterojunction. (From Ref. [93].)

stranded DNA.[94] Photocurrents were observed in the presence of TEOA upon irradiation at 405 nm where the CdS particles absorb. Conduction band electrons from particles near the electrode were directly ejected into the electrode. Electrostatic binding of $Ru(NH_3)_6^{3+}$ to the DNA increased the current. Obviously, the ruthenium complex was active as electron relay (Fig. 11).

Nanocomposite Ni/TiO_2 films were derivatized with hexacyanoferrate to form nickel salt (NHF).[95] Under chopped illumination photocurrent signals indicated photoelectron–hole recombination, mediated by the hexacyanoferrate redox electrolyte, if the potential was below the NHF redox regime. At higher potentials the oxidized NHF centers accepted electrons from the excited TiO_2 with enhanced quantum yield. If a nickel electrode was derivatized with hexacyanoferrate and TiO_2, the electrode exhibited bipolar photoactivity: the photocurrent switched from cathodic to anodic. Below the NHF potential the photogenerated holes from TiO_2 were transferred to reduced NHF sites; above the NHF potential, the electrons were transferred to oxidized NHF sites as in the composite film. This result underlines the importance of the microenvironment (Fig. 12).

Hydrogen Production

After a burst of articles in the early 1980s, little work has been done on hydrogen production recently. TiO_2 and

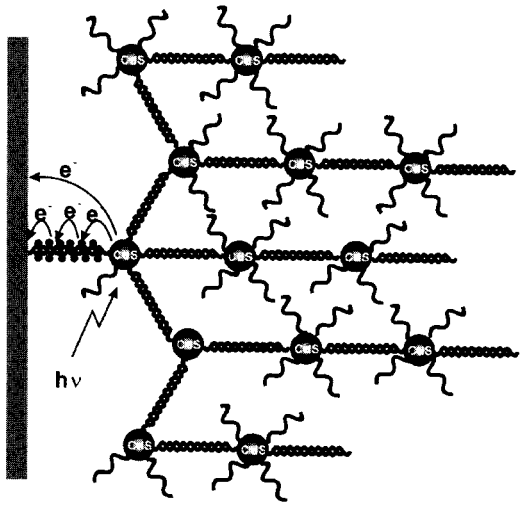


Fig. 11 Model of oligonucleotide-DNA linked arrays of CdS nanoparticles. (From Ref. [94].)

Nb_2O_5 nanoparticles have been used in integrated systems for light-induced vectorial electron transfer and hydrogen production.[96] The electron transfer system consisted of an electron donor, Ruthenium-bipyridyl-dicarboxylate as a sensitizer, the semiconductor as initial electron acceptor, and a viologen as electron relay to the platinum catalyst. This system was spatially organized by a linear channel zeolite (zeolite L or mordenite), which was internally platinized. The semiconducting quantum particles were grown by hydrolysis within the channels and on the external surface of the zeolites. The channel openings (0.71- and 0.65-nm diameter) did not allow the entrance of the ruthenium complex, which was adsorbed exclusively on the surface of the external particles. The viologens were able to enter the channels. In this way the electron donors were separated from the hydrogen catalyst. Despite very efficient initial charge separation no hydrogen (TiO_2 system) or hydrogen with only 0.01% quantum yield (Nb_2O_5) was evolved in the presence of a reversible electron donor (e.g., J^-). The quantum yield was, however, 1% at 450 nm for the system zeolite/TiO_2/MV^{2+}/RuL_3^{2+} in the presence of sacrificial electron donors such as TEOA or EDTA. It should be noted that the initial charge injection from the Ru complex into the semiconductor particle is essential for the success. Without the particles with self-assembly of the MV^{2+} and $Ru(Bipy)_3^{2+}$ onto/into the zeolites very little hydrogen was produced. On the other hand, the H_2-evolution rate was highest in the case of direct bandgap excitation (Fig. 13).

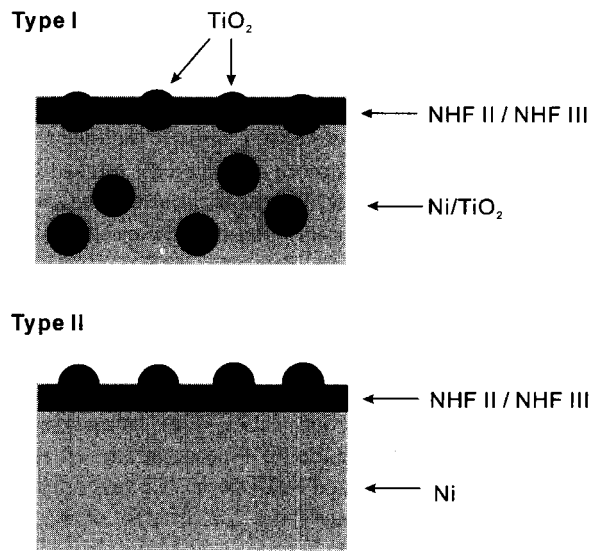


Fig. 12 Model of nickel hexacyanoferrate (NHF) films on TiO_2 or nickel electrodes. Type I was photoactive above an applied potential of 0.4 V. Type II switches from cathodic to anodic photobehavior upon change of potential or Ni(II)/Ni(III) ratio.

SILICA PARTICLES

Silica particles are usually prepared by the method of Stöber et al.[97] which is the polycondensation of tetraethoxysilane (TES). It leads to a dense SiO_2

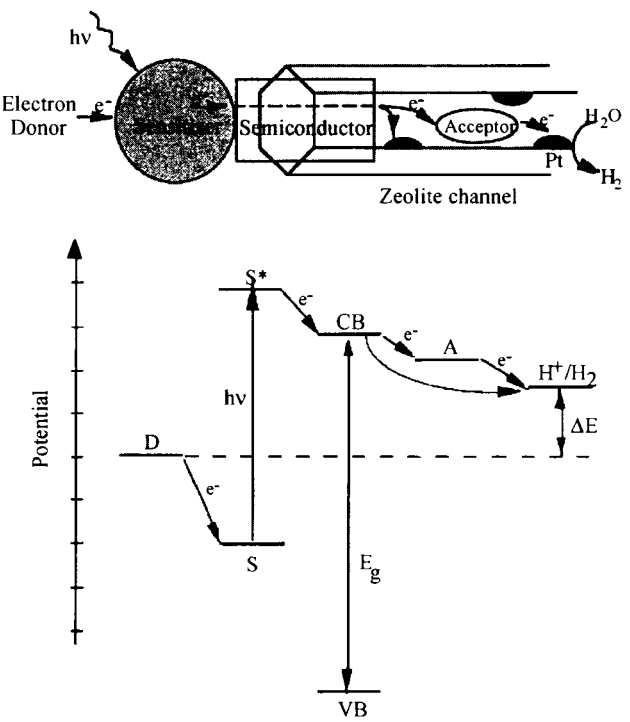


Fig. 13 Model of a sensitized semiconductor–zeolite system for light-induced hydrogen production. (From Ref. [96].)

network, if the conversion of SiOEt groups and the cross-linking are perfect. Otherwise, the hydrophilic and charge stabilized particles contain nanosized pores. The negative surface charge has been used to stabilize the products of photo-induced charge separation.[98] More recently, interest has grown in concentrated dispersions of the silica particles. For this purpose sterically stabilized core-shell particles have been synthesized by surface esterification with octadecanol.[99] These organosilica particles are dispersed in organic solvents and several different colloid phases have been produced: colloidal fluids with only short-range positional order, colloidal glasses with a frozen-in fluid structure, colloidal crystals with long-range positional correlations, and thermotropic gels. These systems are of special interest, when single particles become visual by covalently bound fluorescent dyes. There are three possible positions for the dye: on the particle surface (type a), in a thin shell in the particle interior (type b), and distributed in the volume of an inner core (type c). All three types have been realized by introducing the dye at different stages of the particle synthesis. Isocyanate derivatives of fluoresscein, coumarin, azobenzene, and pyren have been used to be covalently bound to the silane coupling agent 3-aminopropyl-triethoxysilane (APS).[100] Type c particles[101,102] were applied in studies by fluorescence confocal scanning laser microscopy (FCSLM), type b particles[103] in fluorescence recovery after photobleaching (FRAP) experiments. Apart from these investigations centered around the question of interparticle structure and particle diffusion, photochemical studies have not been undertaken. A new method of preparing dye-labeled particles has recently been published.[104,105] Trimethoxymethylsilanes (TMOMS) were used for the synthesis together with chlorobenzyl-trimethoxysilane, leading to lower cross-link density and higher porosity as compared to the particles from TES. Thus carboxylate derivatives of the dyes were able to diffuse through the micropores and were bound by esterification with the chlorobenzyl group. For particles to be stable in aqueous medium, the coating has to be carried out with bifunctional compounds, one group being a terminal amino group.

Fuhrhop[106] started the Stöber synthesis with silicon tetrapropoxide and aminated the surface with a silylchloride derivative. Judging from TEM pictures the particles of at least 60-nm diameter had a smooth surface. In a further derivatization step gaps in monolayers, as already realized on gold particles, were constructed. Tetracarboxyphenylporphyrin was covalently bound to the surface through four amide bonds. Polyoxyethylene as terminal group rendered the particles especially hydrophilic. In contrast to the case of porphyrin bound to gold particles the photophysical properties of the porphyrin were similar to that of corresponding porphyrins in solution (e.g., the fluorescence yield). The triplet state was detected by absorption changes around 780 nm with a relaxation time larger than in solution. Negatively charged *o*-naphthoquinone-sulfonate was able to penetrate to the porphyrine and to accelerate the triplet relaxation by electron transfer (the singlet state was also affected) (Fig. 14).

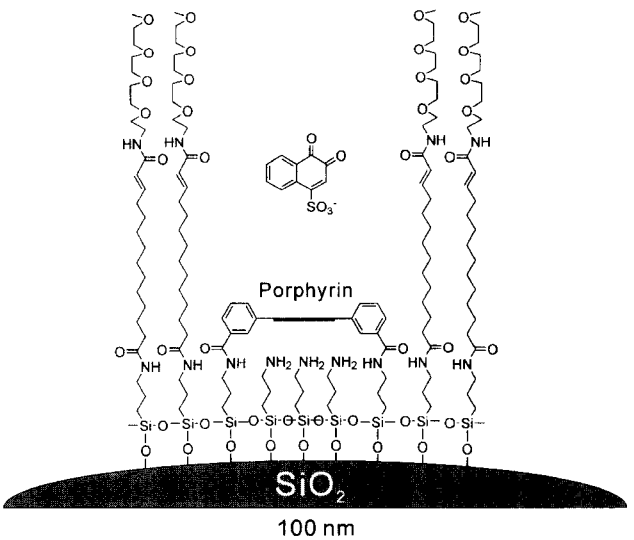


Fig. 14 Model of a porphyrin-based membrane gap. (Fuhrhop, unpublished.)

Silica particles are commercially available in different qualities. Cab-O-Sil, which has been sintered at 500°C, is the least porous material and best suited for the attachment of monolayers. Mallouk[107] has formed multilayers by alternating assembly of α,ω-bis-phosphonic acid-bolaamphiphiles and Zr(IV)-salts on Cab-O-Sil particles terminated by phosphonic acid.

CONCLUSION

Membrane coating of nanoparticles enables to vary the properties and interactions of individual particles, thereby also determining their stability in a given solvent. The ordering of these membranes is less than that of SAMs on planar surfaces and therefore a subject of intense research. Photochemical and photophysical reactions of dyes incorporated into the membranes help to elucidate the order and other properties of the membranous coatings, especially on gold particles. On silver and gold particles, surface-enhanced reactions such as Raman and resonance Raman scattering are analytical tools. Fluorescent dyes on silica particles are used to investigate the structure of concentrated colloidal solutions. On semiconductor particles, dye coatings have a more practical purpose: sensitization for the absorption of visible light makes the

particles interesting in the context of solar energy conversion. The construction of photoelectrochemical cells, the production of hydrogen, and the photochemical detoxification of waste materials, especially in water, are projects that will still become more important in the future. The general trend is to leave the single-particle systems and come to the design of assemblies of nanoparticles with a well-defined architecture. The particles are cross-linked by membranes or polymers, provided with dyes or other molecules, to realize defined distances and surfaces with special properties and finally to arrive at new complex composite materials. Until now for photochemical reactions the large surface of the particles has been the decisive advantage over planar surfaces. The size dependence of optical, electrochemical, and thermodynamic properties of quantum particles has scarcely been exploited and is surely more essential for the design of special optical materials for nonlinear optics and of nanoelectronic circuits.

ACKNOWLEDGMENT

Financial support by the Deutsche Forschungsgemeinschaft (SFB 348 "Mesoscopic Systems"), the European TMR research network "Carbohydrate Recognition," the Fonds der Deutschen Chemischen Industrie and by the FNK of the Free University is gratefully acknowledged.

REFERENCES

1. Turkevich, J.; Stevenson, P.C.; Hiller, J. A study of the nucleation and growth processes in the synthesis of colloidal gold. Discuss. Faraday Soc. **1951**, *11*, 55–75.
2. Cassagneau, T.; Fendler, J.H. Preparation and layer-by-layer self-assembly of silver nanoparticles capped by graphite oxide nanosheets. J. Phys. Chem., B **1999**, *103*, 1789–1793.
3. Hostetler, M.J.; Wingate, J.E.; Zhong, C.-J.; Harris, J.E.; Vachet, R.W.; Clark, M.R.; Londono, J.D.; Green, S.J.; Stokes, J.J.; Wignall, G.D.; Glish, G.L.; Porter, M.D.; Evans, N.D.; Murray, R.W. Alkanethiolate gold cluster molecules with core diameters from 1.5 to 5.2 nm: Core and monolayer properties as a function of core size. Langmuir **1998**, *14*, 17–30.
4. Turkevich, J.; Kim, G. Palladium: Preparation and catalytic properties of particles of uniform size. Science **1970**, *169*, 873–879.
5. Rampino, L.D.; Nord, F.F. Preparation of palladium and platinum synthetic high polymer catalysts and the relationship between particle size and rate of hydrogenation. J. Am. Chem. Soc. **1941**, *63*, 2745–2749.
6. Brust, M.; Walker, M.; Bethell, D.; Schiffrin, D.J.; Whyman, R. Synthesis of thiol-derivatised gold nanoparticles in a two-phase liquid–liquid system. J. Chem. Soc., Chem. Commun. **1994**, 801–802.
7. Hostetler, M.J.; Green, S.J.; Stokes, J.J.; Muray, R.W.J. Monolayers in three dimensions: Synthesis and electrochemistry of ω-functionalized alkanethiolate-stabilized gold cluster compounds. J. Am. Chem. Soc. **1996**, *118*, 4212–4213.
8. Reetz, M.T.; Helbig, W. Size-selective synthesis of nanostructured transition metal clusters. J. Am. Chem. Soc. **1994**, *116*, 7401–7402.
9. Han, M.Y.; Queck, C.H. Photochemical synthesis in formamide and room-temperature coulomb staircase behavior of size-controlled gold nanoparticles. Langmuir **2000**, *16*, 362–367.
10. Link, S.; El-Sayed, M.A. Size and temperature dependence of the plasmon absorption of colloidal gold nanoparticles. J. Phys. Chem., B **1999**, *103*, 4212–4217.
11. Quinten, M.; Kreibig, U. Optical properties of aggregates of small metal particles. Surf. Sci. **1986**, *172*, 557–577.
12. Rozsnyai, L.F.; Wrighton, M.S. Selective electrochemical deposition of polyaniline via photopatterning of a monolayer modified substrate. J. Am. Chem. Soc. **1994**, *116*, 5993–5994.
13. Wollman, E.W.; Kang, D.; Frisbie, C.D.; Lorkovic, I.M.; Wrighton, M.S. Photosensitive self-assembled monolayers on gold: Photochemistry of surface-C azide and cyclopentadienylmanganese tricarbonyl. J. Am. Chem. Soc. **1994**, *116*, 4395–4404.
14. Jocys, G.J.; Workentin, M.S. α-Diazo ketone self-assembled monolayer modified electrode: A proposed photoreactive template for electrode derivatisation. Chem. Commun. **1999**, 839–840.
15. Pitters, J.L.; Kovar, M.; Griffiths. K.; Norton, P.R.; Workentin, M.S. Reactive intermediates on metal surfaces: A ketene monolayer on single crystal platinum generated by photolysis of pyridyl α-diazoketones. Angew. Chem., Int. Ed. **2000**, *39*, 2144–2147.
16. Wolf, M.; Fox, M.A. Photochemistry and surface properties of self-assembled monolayers of *cis*- and cyano-4′-(10-thiodecoxy)stilbene on polycrystalline gold. J. Am. Chem. Soc. **1995**, *117*, 1845–1846.
17. Willner, I.; Doron, A.; Katz, E. Gated molecular and biomolecular optoelectronic systems via photoisomerizable monolayer electrodes. J. Phys. Org. Chem. **1998**, *11*, 546–560.

18. Wolf, M.O.; Fox, M.A. Photoisomerization and photodimerization in self-assembled monolayers of *cis*- and *trans*-4-cyano-4′-(10-mercapatodecoxy)-stilbene on gold. Langmuir **1996**, *12*, 955–962.
19. Kim, T.; Chan, K.C.; Crooks, R.M. Polymeric self-assembled monolayers: 4. Chemical, electrochemical, and thermal stability of ω-functionalized, self-assembled diacetylenic and polydiacetylenic monolayers. J. Am. Chem. Soc. **1997**, *119*, 189–193.
20. Fox, M.A.; Wooten, M.D. Characterization, adsorption, and photochemistry of self-assembled monolayers of 10-thiodecyl 2-anthryl ether on gold. Langmuir **1997**, *13*, 7099–7105.
21. Li, W.; Lynch, V.; Thompson, H.; Fox, M.A. Self-assembled monolayers of 7-(10-thiodecoxy)coumarin on gold: Synthesis, characterization, and photodimerization. J. Am. Chem. Soc. **1997**, *119*, 7211–7217.
22. Imahori, H.; Arimura, M.; Hanada, T.; Nishimura, Y.; Yamazaki, I.; Sakata, Y.; Fukuzumi, S.J. Photoactive three-dimensional monolayers: Porphyrin-alkanethiolate-stabilized gold clusters. Am. Chem. Soc. **2001**, *123*, 335–336.
23. Aguila, A.; Murray, R.W. Monolayer-protected clusters with fluorescent dansyl ligands. Langmuir **2000**, *16*, 5949–5954.
24. Ruzgas, L.L.; Gaigalas, A.K. Fluorescence from Alexa 488 fluorophore immobilized on a modified gold electrode. Langmuir **1999**, *15*, 6358–6363.
25. Chen, M.M.Y.; Katz, A. Steady-state fluorescence-based investigation of the interaction between protected thiols and gold nanoparticles. Langmuir **2002**, *18*, 2413–2420.
26. Wang, T.; Zhang, D.; Xu, W.; Yang, J.; Han, R.; Zhu, D.Preparation, characterization, and photophysical properties of alkanethiols with pyrene units-capped gold nanoparticles: unusual fluorescence enhancement for the aged solutions of these gold nanoparticles. Langmuir **2002**, *18*, 1840–1848.
27. Zhang, J.; Whitesell, J.K.; Fox, M.A. Photoreactivity of self-assembled monolayers of azobenzene or stilbene derivatives capped on colloidal gold clusters. Chem. Mater. **2001**, *13*, 2323–2331.
28. Hu, J.; Zhang, J.; Liu, F.; Kittredge, K.; Whitesell, J.K.; Fox, M.A. Competitive photochemical reactivity in a self-assembled monolayer on a colloidal gold cluster. J. Am. Chem. Soc. **2001**, *123*, 1464–1470.
29. Evans, S.D.; Johnson, S.R.; Ringsdorf, H.; Williams, L.M.; Wolf, H. Photoswitching of azobenzene derivatives formed on planar and colloidal gold surfaces. Langmuir **1998**, *14*, 6436–6440.
30. Kell, A.J.; Stringle, D.L.B.; Workentin, M.S. Norrish type II photochemical reaction of an aryl ketone on a monolayer-protected gold nanocluster. Development of a probe of conformational mobility. Org. Lett. **2000**, *2* (21), 3381–3384.
31. Thomas, K.G.; Kamat, P.V. Making gold nanoparticles glow: Enhanced emission from a surface-bound fluoroprobe. J. Am. Chem. Soc. **2000**, *122*, 2655–2656.
32. Hone, D.C.; Walker, P.I.; Evans-Gowing, R.; FitzGerald, S.; Beeby, A.; Chambrier, I.; Cook, M.J.; Russell, D.A. Generation of cytotoxic singlet oxygen via phthalocyanine-stabilized gold nanoparticles: A potential delivery vehicle for photodynamic therapy. Langmuir **2002**, *18*, 2985–2987.
33. Kometani, N.; Tsubonishi, M.; Fujita, T.; Asami, K.; Yonezawa, Y. Preparation and optical absorption spectra of dye-coated Au, Ag, and Au/Ag colloidal nanoparticles in aqueous solutions and in alternate assemblies. Langmuir **2001**, *17*, 578–580.
34. Hranisavljevic, J.; Dimitrijevic, N.M.; Wurtz, G.A.; Wiederrecht, G.P. Photoinduced charge separation reactions of J-aggregates coated on silver nanoparticles. J. Am. Chem. Soc. **2002**, *124*, 4536–4537.
35. Baschong, W.; Lucocq, J.M.; Roth, J. Thiocyanate gold: Small (2–3 nm) colloidal gold for affinity cytochemical labelling in electron microscopy. Histochemistry **1985**, *83*, 409–411.
36. Henglein, A. Radiolytic preparation of ultrafine colloidal gold particles in aqueous solution: Optical spectrum, controlled growth, and some chemical reactions. Langmuir **1999**, *14*, 6738–6744.
37. Chandrasekharan, N.; Kamat, P.V.; Hu, J.; Jones, G. Dye-capped gold nanoclusters: Photoinduced morphological changes in gold/rhodamine 6G nanoassemblies. J. Phys. Chem., B **2000**, *104*, 11103–11109.
38. Johnson, C.K.; Soper, J.A. Nonlinear surface-enhanced spectroscopy of silver colloids and pyridine: Hyper-Raman and second-harmonic scattering. J. Phys. Chem. **1989**, *93*, 7281–7285.
39. Clays, K.; Hendricks, E.; Triest, M.; Persoons, A. Second-order nonlinear optics in isotropic liquids: Hyper-Rayleigh scattering in solution. J. Mol. Liq. **1995**, *67*, 133–155.
40. Baranov, A.V.; Inoue, K.; Toba, K.; Yamanaka, A.; Petrov, V.I.; Fedorov, A.V. Resonant hyper-Raman and second harmonic scattering in a CdS quantum dot system. Phys. Rev., B Condens. Matter **1996**, *53*, R1721–R1724.
41. Vance, F.W.; Lemon, V.I.; Hupp, J.T. Enormous hyper-Rayleigh scattering from nanocrystalline gold particle suspensions. J. Phys. Chem., B **1998**, *102*, 10091–10093.
42. Johnson, R.C.; Hupp, J.T. *Metal Nanoparticles. Synthesis Characterization and Applications*;

Feldheim, D., Foss, C., Eds.; Marcel-Dekker: New York, 2000.
43. Shiang, J.J.; Heath, J.R.; Collier, C.P.; Saykally, R.J. Cooperative phenomena in artificial solids made from silver quantum dots: The importance of classical coupling. J. Phys. Chem., B **1998**, *102*, 3425–3430.
44. Chumanov, G.; Sokalov, K.; Gregory, B.; Cotton, T.M. Colloidal metal films as a substrate for surface-enhanced spectroscopy. J. Phys. Chem. **1995**, *99*, 9466–9471.
45. Feilchenfeld, H.; Chumanov, G.; Cotton, T.M. Photoreduction of methylviologen adsorbed on silver. J. Phys. Chem. **1996**, *100*, 4937–4943.
46. Moskovits, M. Surface-enhanced spectroscopy. Rev. Mod. Phys. **1985**, *57*, 783–826.
47. Harfenist, S.A.; Wang, Z.L.; Alvarez, M.M.; Vezmar, I.; Whetten, R.L. Highly oriented molecular Ag nanocrystal arrays. J. Phys. Chem. **1996**, *100*, 13904–13910.
48. Brousseau, L.C., III; Novak, J.P.; Marinakos, S.M.; Feldheim, D.L. Assembly of phenylacetylene-bridged gold nanocluster dimers and trimers. Adv. Mater. **1999**, *11*, 447–449.
49. Novak, J.P.; Feldheim, D.L. Assembly of phenylacetylene-bridged silver and gold nanoparticle arrays. J. Am. Chem. Soc. **2000**, *122*, 3979–3980.
50. Novak, J.P.; Brousseau, L.C., III; Vance, F.W.; Johnson, R.C.; Lemon, B.I.; Hupp, J.T.; Feldheim, D.L. Nonlinear optical properties of molecularly bridged gold nanoparticle arrays. J. Am. Chem. Soc. **2000**, *122*, 12029–12030.
51. Šluofová-Srnová, I.; Vlčková, B. Two-dimensional assembling of Au nanoparticles mediated by tetrapyridylporphine molecules. Nano Lett. **2002**, *2* (2), 121–125.
52. Lahav, M.; Gabriel, T.; Shipway, A.N.; Willner, I. Assembly of a Zn(II)-porphyrin-bipyridinium dyad and Au-nanoparticle superstructures on conductive surfaces. J. Am. Chem. Soc. **1999**, *121*, 258–259.
53. Lahav, M.; Heleg-Shabtai, V.; Wasserman, J.; Katz, E.; Willner, I.; Duerr, H.; Hu, Y.-Z.; Bossmann, S.H. Photoelectrochemistry with integrated photosensitizer-electron acceptor and Au-nanoparticle arrays. J. Am. Chem. Soc. **2000**, *122*, 11480–11487.
54. Sudeep, P.K.; Ipe, B.I.; Thomas, K.G.; George, M.V.; Barazzouk, K.S.; Hotchandani, S.; Kamat, P.V. Fullerene-functionalized gold nanoparticles. A self-assembled photoactive antenna-metal nanocore assembly. Nano Lett. **2002**, *2* (1), 29–35.
55. Kuwahara, Y.; Akiyama, T.; Yamada, S. Facile fabrication of photoelectrochemical assemblies consisting of gold nanoparticles and a *tris*(2,2′-bipyridine)ruthenium(II)-viologen linked thiol. Langmuir **2001**, *17*, 5714–5716.
56. Imahori, H.; Norieda, H.; Yamada, H.; Nishimura, Y.; Yamazaki, I.; Sakata, Y.; Fukuzumi, S. Light-harvesting and photocurrent generation by gold electrodes modified with mixed self-assembled monolayers of boron–dipyrrin and ferrocene–porphyrin–fullerene triad. J. Am. Chem. Soc. **2001**, *123*, 100–110.
57. Li, G.; Fuhrhop, J.H. Anticorrosive lipid monolayers with rigid walls around porphyrin-based 2 nm gaps on 20 nm gold particles. Langmuir **2002**, *18*, 7740–7747.
58. Fudickar, W.; Zimmermann, J.; Ruhlmann, L.; Roeder, B.; Siggel, U.; Fuhrhop, J.-H. Fluorescence quenching and size selective heterodimerization of a porphyrin adsorbed to gold and embedded in rigid membrane gaps. J. Am. Chem. Soc. **1999**, *121*, 9539–9545.
59. Skupin, M.; Li, G.; Fudickar, W.; Zimmermann, J.; Roeder, B.; Fuhrhop, J.-H. Methylammonium groups at the solid walls of nanometer-sized, water-filled monolayer gaps as binding sites for a tetraanionic porphyrin. J. Am. Chem. Soc. **2001**, *123*, 3454–3461.
60. Alivisatos, A.P. Semiconductor clusters, nanocrystals, and quantum dots. Science **1996**, *271*, 933–937.
61. Hickey, S.G.; Rilley, D.J. Photoelectrochemical studies of CdS nanoparticle-modified electrodes. J. Phys. Chem., B **1999**, *103*, 4599–4602.
62. Klein, D.L.; Roth, R.; Kim, A.K.L.; Alivisatos, A.P.; McEuen, P.L. A single-electron transistor made from a cadmiumselenide nanocrystal. Nature **1997**, *389*, 699–701.
63. Meldrum, F.C.; Flath, J.; Knoll, W. Chemical deposition of PbS on self-assembled monolayers of 16-mercaptohexadecanoic acid. Langmuir **1997**, *13*, 2033–2049.
64. Spanhel, L.; Weller, H.; Fojtik, A.; Henglein, A. Photochemistry of semiconductor colloids: 17. Strong luminescing CdS and CdS-Ag_2S particles. Ber. Bunsenges. Phys. Chem. **1987**, *91*, 88–94.
65. Rizza, R.; Fitzmaurice, D.; Hearne, S.; Hughes, G.; Spoto, G.; Ciliberto, E.; Kerp, H.; Schropp, R. Self-assembly of monolayers of semiconductor nanocrysallites. Chem. Mater. **1997**, *9*, 2969–2982.
66. Hotchandani, S.; Kamat, P.V. Charge-transfer processes in coupled semiconductor systems. Photochemistry and photoelectrochemistry of the colloidal CdS–ZnO system. J. Phys. Chem. **1992**, *96*, 6834–6839.
67. Bedja, I.; Hotchandani, S.; Kamat, P.V. Preparation and photoelectrochemical characterization of thin

SnO_2 nanocrystalline semiconductor films and their sensitization with bis(2,2′-bipyridine) (2,2′-bipyridine-4,4′-dicarboxylic acid)ruthenium(II) complex. J. Phys. Chem. **1994**, *98*, 4133–4140.
68. Gerischer, H.; Willig, F. Reaction of excited dye molecules at electrodes. Top. Curr. Chem. **1976**, *61*, 31–84.
69. Nozik, Y.; Memming, R. Physical chemistry of semiconductor–liquid interfaces. J. Phys. Chem. **1996**, *100*, 13061–13078.
70. Lawless, D.; Kapoor, S.; Meisel, D. Bifunctional capping of CdS nanoparticles and bridging to TiO_2. J. Phys. Chem., B **1995**, *99*, 10329–10335.
71. Vogel, R.; Pohl, K.; Weller, H. Sensitization of highly porous, polycrystalline TiO_2 electrodes by quantum sized CdS. Chem. Phys. Lett. **1990**, *174*, 241–246.
72. Kortan, A.R.; Hull, R.; Opila, R.L.; Stams, D.A.; Thomas, T.D.; McLaren, D.C.; Ji, D.; Morton, T.H. Nucleation and growth of CdSe on ZnS quantum crystallite seeds, and vice versa, in inverse micelle media. J. Am. Chem. Soc. **1990**, *112*, 1327–1332.
73. Henglein, A. Photochemistry of colloidal cadmium sulfide: 2. Effects of adsorbed methyl viologen and of colloidal platinum. J. Phys. Chem. **1982**, *86*, 2291–2293.
74. Duonghong, D.; Ramsden, J.; Graetzel, M. Dynamics of interfacial electron-transfer processes in colloidal semiconductor systems. J. Am. Chem. Soc. **1982**, *104*, 2977–2985.
75. Desilvestro, J.; Graetzel, M.; Kavan, L.; Moser, J. Highly efficient sensitization of titanium dioxide. J. Am. Chem. Soc. **1985**, *107*, 2988–2990.
76. Vrachnou, E.; Vlachopoulos, N.; Graetzel, M. Efficient visible light sensitization of TiO_2 by surface complexation with $Fe(CN)_6^{4-}$. J. Chem. Soc., Chem. Commun. **1987**, 868–870.
77. Kalyanasundaram, K.; Vlachopoulos, N.; Krishnan, V.; Monnier, A.; Graetzel, M. Sensitization of TiO_2 in the visible light region using zinc porphyrins. J. Phys. Chem. **1987**, *91*, 2342–2347.
78. Liska, P.; Vlachopoulos, N.; Nazeeruddin, M.K.; Comte, P.; Graetzel, M. *cis*-Diaquabis(2,2′-bipyridyl-4,4′-dicarboxylate)-ruthenium(II) sensitizes wide band gap oxide semiconductors very efficiently over a broad spectral range in the visible. J. Am. Chem. Soc. **1988**, *110*, 3686–3687.
79. Vlachopoulos, N.; Liska, P.; Augustynski, J.; Graetzel, M. Very efficient visible light energy harvesting and conversion by spectral sensitization of high surface area polycrystalline titanium dioxide films. J. Am. Chem. Soc. **1988**, *110*, 1216–1220.
80. Enea, O.; Moser, J.; Graetzel, M. Achievement of incident photon to electric current conversion yields exceeding 80% in the spectral sensitization of titanium dioxide by coumarin. J. Electroanal. Chem. **1989**, *259*, 59–65.
81. Kay, A.; Humphry-Baker, R.; Graetzel, M. Artificial photosynthesis: 2. Investigations on the mechanism of photosensitization of nanocrystalline TiO_2 solar cells by chlorophyll derivatives. J. Phys. Chem. **1994**, *98*, 952–959.
82. Hilgendorff, M.; Sundstroem, V. Dynamics of electron injection and recombination of dye-sensitized TiO_2 particles. J. Phys. Chem., B **1998**, *102*, 10505–10514.
83. Asbury, J.B.; Hao, E.; Wang, Y.; Lian, T. Bridge length-dependent ultrafast electron transfer from re polypyridyl complexes to nanocrystalline TiO_2 thin films studied by femtosecond infrared spectroscopy. J. Phys. Chem. **2000**, *104*, 11957–11964.
84. Asbury, J.B.; Wang, Y.Q.; Hao, E.; Gosh, H.N.; Lian, T. Evidence of hot excited state electron injection from sensitizer molecules to TiO_2 nanocrystalline thin films. Res. Chem. Intermed. **2001**, *27*, 393.
85. Bonhote, P.; Moser, J.E.; Humphry-Baker, R.; Vlachopoulos, N.; Zakeeruddin, S.M.; Walder, L.; Grätzel, M. Long-lived photoinduced charge separation and redox-type photochromism on mesoporous oxide films sensitized by molecular dyads. J. Am. Chem. Soc. **1999**, *121*, 1324–1336.
86. Ramakrishna, G.; Ghosh, H.N. Efficient electron injection from twisted intramolecular charge transfer (TICT) state of 7-diethyl amino coumarin 3-carboxylic acid (D-1421) dye to TiO_2 nanoparticle. J. Phys. Chem., A **2002**, *106*, 2545–2553.
87. Grätzel, M. Nanocrystalline Electronic Junctions. In *Semiconductor Nanoclusters—Physical Chemical and Catalytic Aspects*; Kamat, P.V., Meisel, D., Eds.; Elsevier: Amsterdam, 1997; 335.
88. Khazraji, A.C.; Hotchandani, S.; Das, S.; Kamat, P.V. Controlling dye (merocyanine-540) aggregation on nanostructured TiO_2 films. An organized assembly approach for enhancing the efficiency of photosensitization. J. Phys. Chem., B **1999**, *103*, 4693–4700.
89. Kay, A.; Graetzel, M. Artificial photosynthesis: 1. Photosensitization of TiO_2 solar cells with chlorophyll derivatives and related natural porphyrins. J. Phys. Chem. **1993**, *97*, 6272–6277.
90. O'Regan, B.; Graetzel, M. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO_2 films. Nature **1991**, *353*, 737–739.
91. O'Regan, B.; Schwarz, D.T. Large enhancement in photocurrent efficiency caused by UV illumination

of the dye-sensitized heterojunction TiO_2/Ru-LL'NCS/CuSCN: Initiation and potential mechanisms. Chem. Mater. **1998**, *10*, 1501–1509.
92. Hagen, G.; Haarer, D.; Schaffrath, W.; Otschik, P.; Fink, R.; Bacher, A.; Schmidt, H.-W. Novel hybrid solar cells consisting of inorganic nanoparticles and an organic hole transport material. Synth. Met. **1997**, *89*, 215–220.
93. Bach, U.; Lupo, D.; Comte, P.; Moser, J.E.; Weissoertel, F.; Salbeck, J.; Spreitzer, H.; Graetzel, M. Solid-state dye-sensitized mesoporous TiO_2 solar cells with high photon-to-electron conversion efficiencies. Nature **1998**, *395* (8), 583–585.
94. Willner, I.; Patolsky, F.; Wasserman, J. Photoelectrochemistry with controlled DNA-cross-linked CdS nanoparticle arrays. Angew. Chem. **2001**, *113* (10), 1913–1916.
95. de Tacconi, N.R.; Carmona, J.; Rejeshwar, K. Chemically modified Ni/TiO_2 nanocomposite films: Charge transfer from photoexcited TiO_2 particles to hexacyanoferrate redox centers within the film and unusual photoelectrochemical behavior. J. Phys. Chem., B **1997**, *101* (49), 10151–10154.
96. Kim, Y.I.; Keller, S.W.; Krueger, J.S.; Yonemoto, E.H.; Saupe, G.B.; Mallouk, T.E. Photochemical charge transfer and hydrogen evolution mediated by oxide semiconductor particles in zeolite-based molecular assemblies. J. Phys. Chem., B **1997**, *101*, 2491–2500.
97. Stöber, W.; Fink, A.; Bohn, E. Controlled growth of monodisperse silica spheres in the micron size range. J. Colloid Interface Sci. **1968**, *26*, 62–69.
98. Willner, I.; Yango, J.-M.; Laane, C.; Otvos, J.W.; Calvin, M. The function of SiO_2 colloids in photoinduced redox reactions. Interfacial effects on the quenching, charge separation, and quantum yields. J. Phys. Chem. **1981**, *85*, 3277–3282.
99. van Helden, A.K.; Jansen, J.W.; Vrij, A. Preparation and characterization of spherical monodisperse silica dispersions in nonaqueous solvents. J. Colloid Interface Sci. **1981**, *81* (2), 354–368.
100. van Blaaderen, A.; Vrij, A. Synthesis and characterization of colloidal dispersions of fluorescent, monodisperse silica spheres. Langmuir **1992**, *8*, 2921–2931.
101. van Blaaderen, A.; Imhof, A.; Hage, W.; Vrij, A. Three-dimensional imaging of submicrometer colloidal particles in concentrated suspensions using confocal scanning laser microscopy. Langmuir **1992**, *8*, 1514–1517.
102. Verhaegh, N.A.M.; van Blaaderen, A. Dispersions of rhodamine-labeled silica spheres: Synthesis, characterization, and fluorescence confocal scanning laser microscopy. Langmuir **1994**, *10*, 1427–1438.
103. van Blaaderen, A.; Peetermans, J.; Maret, G.; Dhont, J.K.G. Long-time self-diffusion of spherical colloidal particles measured by fluorescence recovery after photobleaching. J. Chem. Phys. **1992**, *96*, 4591–4603.
104. Graf, C.; Schaertl, W.; Maskos, M.; Schmidt, M. Tracer diffusion of polyorganosiloxane nanoparticles in solution: Effects of tracer topology and particle concentration. J. Chem. Phys. **2000**, *112*, 3031–3039.
105. Graf, C.; Schaertl, W.; Fischer, K.; Hugenberg, N.; Schmidt, M. Dye-labeled poly(organosiloxane) microgels with core-shell architecture. Langmuir **1999**, *15*, 6170–6180.
106. Li, G.; Bhosale, S.V.; Wang, T.; Hackbarth, S.; Roeder, B.; Siggel, U.; Fuhrhop. J.-H. Nanowells on silica particles in water containing long-distance porphyrin heterodimers. J. Am. Chem. Soc. **2003**, *125*, 10693–10702.
107. Yang, H.C.; Aoki, K.; Hong, H.-G.; Sackett, D.D.; Arendt, M.F.; Yau, S.-L.; Bell, C.M.; Mallouk, T.E. Growth and characterization of metal(II) alkane-bis-phosphate multilayer thin films on gold surfaces. J. Am. Chem. Soc. **1993**, *115*, 11855–11862.

Photonic Crystal Fibers

P. St.J. Russell
J. C. Knight
T. A. Birks
University of Bath, Bath, United Kingdom

P. J. Roberts
BlazePhotonics Ltd., Bath, United Kingdom

INTRODUCTION

In 1842, the Swiss physicist Daniel Colladon[1] demonstrated that light could be guided along a curved path within a tube of transparent material. Colladon's "light pipe" was simply a narrow stream of water issuing through a hole in the side of a tank. The underlying guidance mechanism in this case was, of course, total internal reflection (TIR)—the phenomenon by which standard single-mode fiber (SMF) guides light. A typical SMF designed for communications wavelengths (Fig. 1a) has a Ge-doped silica core (synthesized by modified chemical vapor deposition),[2] a core-cladding refractive index difference of a few percent, and a core diameter of ~ 10 μm, and possesses astonishing optical clarity (0.2 dB/km at 1550 nm). Although SMF has been outstandingly successful in optical telecommunications, there is still pressure to further improve its performance. For example, optical nonlinearities in the solid glass core are proving complex to understand and awkward to control, and are contributing to bit error rate deterioration over long spans in wavelength division multiplexed systems. In other fields, there is a long-standing need for fibers that can carry higher power; act as versatile sensors or hosts for rare earth ions; and have multiple cores, higher nonlinearities, lower nonlinearities, higher birefringence, and widely engineerable dispersion.

Annular Fibers

In the early 1970s, another kind of structure had been proposed, in theory, as an alternative to SMF (Fig. 1b). This was a cylindrical fiber that guided light by Bragg scattering at a periodic array of annular rings of high and low refractive index arranged around a central core.[3] A team in France has recently succeeded in making a solid-core version of this structure using modified chemical vapor deposition (MCVD).[4] Work is also underway to realize a hollow-core version of a similar structure.[5]

Photonic Crystal Fibers

In 1991, the idea emerged[6] that light might be trapped in a hollow core by means of a two-dimensional "photonic crystal" of microscopic air capillaries running along the entire length of a fiber. Appropriately designed, this array would support a photonic band gap for incidence from air, preventing the escape of light from a hollow core into the cladding and avoiding the need for TIR. It was initially far from obvious that so many new applications and developments would emerge from the photonic crystal fiber (PCF) concept.[7–9] The two basic types—with hollow and solid cores—are illustrated in Fig. 1c and d. In this article, we review its fabrication and guidance mechanisms, numerical modelling techniques, and the numerous and growing number of applications.

Nomenclature

The PCF idea came from the field of photonic band gap materials, not from within fiber optics. For this reason, we prefer the name PCF. It is also sometimes referred to as "holey" fiber (P. R.'s first jocular name for it, before anyone knew whether it could be made) or "microstructure" fiber.

FABRICATION TECHNIQUES

The first working PCF emerged from the drawing tower in late 1995.[10] Since then, structures with increasing perfection and uniformity have been produced in many laboratories, using several different techniques (Fig. 2). The first stage is the production of a "preform"—a macroscopic facsimile of the nanoscopic structure in the drawn PCF. There are many ways to do this, including stacking of capillaries and rods,[11] extrusion,[12–14] sol–gel casting, injection molding, and drilling.[15]

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009350

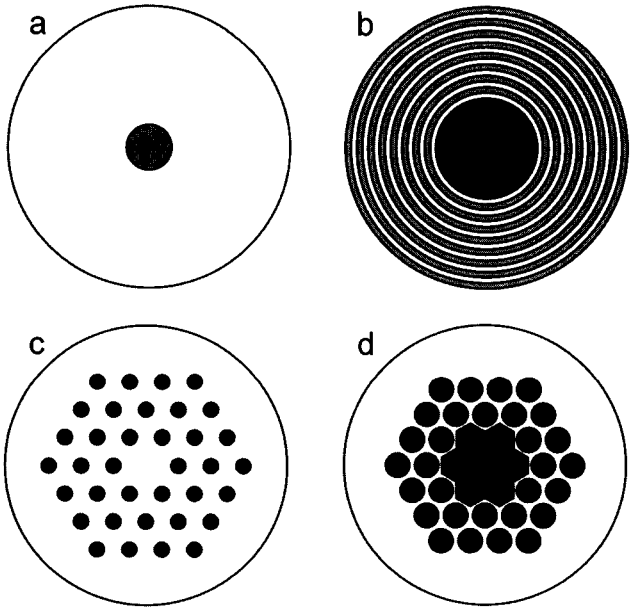


Fig. 1 (a) Standard SMF: outer diameter, 125 μm; core diameter, ~9 μm; core refractive index, a few percent higher than the silica-cladding index (1.46). (b) Annular Bragg fiber. (c) Solid-core PCF. (d) Hollow-core PCF. The black regions are hollow, and the white and grey regions denote pure and doped glass.

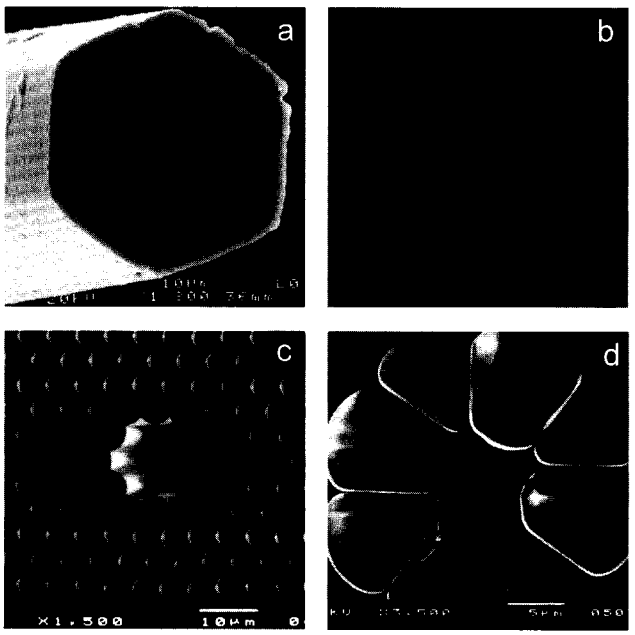


Fig. 2 (a) The first working PCF—the solid glass core is surrounded by a triangular array of 300-nm-diameter air channels, spaced 2.3 μm apart. (From Ref. [11].) (b) Detail of a recent low-loss solid-core PCF (interhole spacing, ~2 μm). (From Refs. [26] and [35].) (c) The first hollow-core PCF. (From Ref. [19].) (d) A small-core PCF extruded from Schott SF6 glass. (From Ref. [13].)

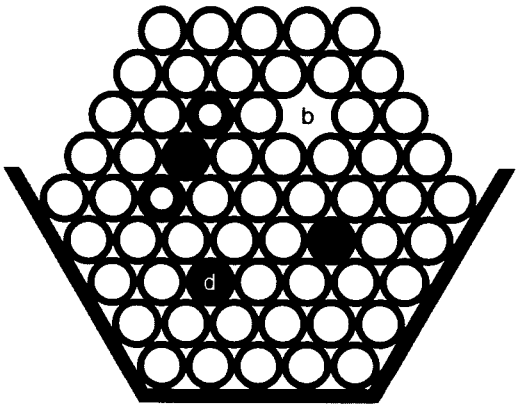


Fig. 3 Preform stack containing (a) birefringent solid core; (b) hollow core; (c) solid isotropic core; and (d) doped core. The capillary diameters are ~1 mm.

The most widely used technique is capillary stacking (Fig. 3), a method that is especially suited to silica glass because of the ready availability of high-quality starting tubes of appropriate dimensions. Typically, meter-length capillaries with an outer diameter of around 1 mm are drawn from a starting tube of high-purity synthetic silica with a diameter of perhaps 20 mm and a length of 1 m. The inner diameter of the capillaries is largely set by the inner/outer diameter ratio of the starting tube, which might lie in the range from 0.3 up to beyond 0.9. The uniformity in diameter of the capillaries (and their circularity) should be controlled to at least 1% of the diameter, and preferably far better. When a suitable quantity of uniform capillaries is available, they are stacked horizontally by hand in a suitable jig to form a close-packed arrangement. The final stack is loosely bound with wire, before being inserted into a snug jacketing tube. It is then mounted in the preform feed assembly for drawing down to fiber.

Extrusion is an attractive alternative route to making PCF from bulk glass. It enables the formation of structures that are not readily achievable by stacking, and requires a minimal amount of processing. Perhaps its most important application is to nonsilica glass PCFs and other materials. Fig. 2d shows the cross section of a fiber extruded, through a metal die, from a commercially available glass (Schott SF6).[13]

CHARACTERISTICS OF PHOTONIC CRYSTAL CLADDING

The simplest photonic crystal cladding is a biaxially periodic, defect-free, composite material with its own well-defined dispersion and band structure. A good knowledge of its properties is essential for understanding the behavior

of the guided modes that form at cores (or "structural defects" in the parlance of photonic crystals). A useful graphical tool is the propagation diagram—a map of the ranges of frequency and axial wavevector component where light is evanescent in all transverse directions regardless of its polarization state (Fig. 4).[16] The vertical axis is the normalized frequency $k\Lambda = \omega\Lambda/c$ (ω is the angular frequency, Λ is the interhole spacing, and c is the velocity of light in vacuum), and the horizontal axis is the normalized axial wavevector $\beta\Lambda$. Light is free to propagate in the white areas and is evanescent (due either to TIR or photonic band gaps) in the black regions. In any subregion of isotropic material (glass or air) at fixed optical frequency, the maximum possible value of $\beta\Lambda$ is given by $k\Lambda n$, where n is the refractive index (at that frequency) of the region under consideration. For $\beta < kn$, light is free to propagate; for $\beta > kn$, it is evanescent; and at $\beta = kn$, the critical angle is reached—denoting the onset of TIR for light incident from a medium of index larger than n. Moving from left to right, the slanted guidelines (included in Fig. 4) denote the transitions from propagation to evanescence for glass, air, and the photonic crystal. The maximum value of $\beta\Lambda$ in the photonic crystal cladding lies in the range $k < \beta\Lambda < kn_s$ (where n_s is the index of silica) as expected of a glass/air composite material. Its maximum axial refractive index depends strongly on frequency, even though neither the air nor the glass is assumed to be dispersive in the analysis. Microstructuring itself creates dispersion, through a balance between transverse energy storage and energy flow that is highly dependent on frequency. As the wavelength of the light falls, the optical fields are better able to distinguish between the glass regions and the air. The light piles up more and more in the glass or the air, causing the effective refractive index "seen" by it to change. In the limit of small wavelength $\Lambda/\lambda \to \infty$, the light is strongly excluded from the air holes by TIR, and the field profile freezes into a shape independent of wavelength. Treating each strand as a metallic waveguide of radius $a = \Lambda/\sqrt{3} - d/2$ and making the approximation that the field amplitude is $J_0(rz_{01}/a)$ (where z_{01} is the first zero of the Bessel function J_0), it may be shown that the maximum axial refractive index in the cladding is:

$$n_{\mathrm{FSM}} \approx \sqrt{n_s^2 - \frac{z_{01}^2\lambda^2}{4\pi^2 a^2}} \qquad (1)$$

in the short wavelength limit. For a more rigorous account of this phenomenon, see Ref. [17].

The finger-shaped regions on the diagram denote full two-dimensional photonic band gaps. Some of these extend into the region $\beta < k$, where light is free to propagate in vacuum, confirming the feasibility of trapping light within an empty microtube.

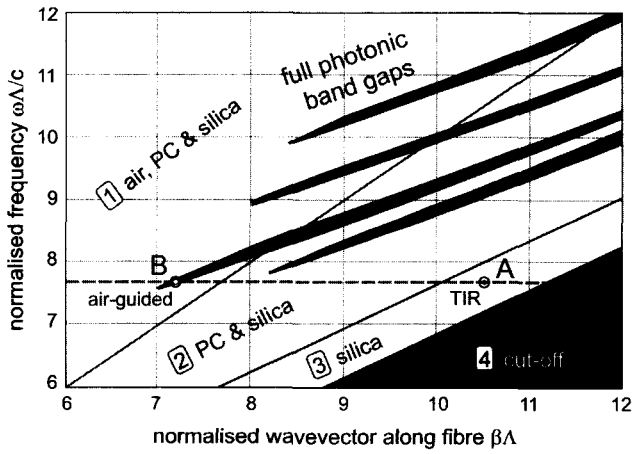


Fig. 4 Propagation diagram for a PCF with 45% air-filling fraction. Note the different regions where light is (4) cut off completely, (3) able to propagate only in silica glass, (2) able to propagate also in the photonic crystal cladding, and (1) able to propagate in all regions. The "fingers" indicate the positions of full two-dimensional photonic band gaps. (From Ref. [16].)

CLASSES OF GUIDANCE

In SMF, guided modes form in the range of axial refractive indices $n_{cl} < n_z < n_{core}$, when light is evanescent in the cladding ($n_z = \beta/k$). In PCF, four distinct guidance mechanisms exist: a modified form of TIR,[17,18] photonic band gap guidance,[19] a frustrated tunneling process (leading to the formation of surface states around the edge of the core),[20] and a low-leakage mechanism based on creating a low density of states in the cladding.[21]

Modified Total Internal Reflection

This type of guidance was observed in late 1995, in the first working PCF, at a central solid core in an array of ~300-nm-diameter air holes, spaced 2.3 μm apart (Fig. 2a). The striking feature of this fiber was that it was "endlessly single-mode" (i.e., the core did not ever seem to become multimode in the experiments, no matter how short the wavelength of the light).[17] Although the guidance in some respects closely resembled conventional TIR, it turned out to have some interesting and unique features that distinguish it markedly from the standard case. These are because of the piecewise discontinuous nature of the core boundary—sections where air holes strongly block the escape of light interspersed with regions of barrier-free glass. In fact, the cladding operates in the Mie scattering regime where transverse effective wavelength (in silica) is comparable with geometrical

substructures in the cladding. The zone of operation in Fig. 4 is $n_{FSM} < n_z < n_{silica}$.

Photonic Band Gap: Hollow Core

Larger air-filling fractions and small interhole spacings are necessary to achieve photonic band gaps in the region $n_z < 1$. The relevant operating region in Fig. 4 is to the left of the vacuum line inside one of the band gap fingers. These conditions ensure that light is free to propagate—and form guided modes—within the hollow core while unable to escape into the cladding. The number N of such modes is controlled by the depth and width of the refractive index "potential well" and is given by:

$$N \approx 2\pi^2\rho^2(n_{high}^2 - n_{low}^2)/\lambda^2 \quad (2)$$

where n_{high} and n_{low} are the refractive indices at the edges of the photonic band gap, ρ is the core radius, and λ is the wavelength of light. Because the band gaps are quite narrow ($n_{high}^2 - n_{low}^2$ is typically a few percent), the hollow core must be large if a guided mode is to exist at all. In the first working hollow-core PCF reported in 1999,[19] the core was formed by omitting seven capillaries from the preform stack (Fig. 2c). An electron micrograph of a more recent hollow-core PCF is shown in Fig. 5.[22]

Photonic Band Gap: Surface States

To observe unambiguous photonic band gap guidance, it is necessary to ensure that the core refractive index is lower than the cladding index. The first such PCF, demonstrated in 1998, consisted of a lattice of air holes arranged in the same way as the carbon rings in graphite. A core was formed by introducing an extra hole at the center of one of the rings, its low index precluding the possibility of TIR guidance.[20] When white light was launched into the core region, a brightly colored mode was transmitted—the colors being dependent on the absolute size to which the fiber was drawn. The modal patterns had six equally strong lobes, disposed in a flowerlike pattern around the central hole. Closer examination revealed that the light was guided not in the air holes but in the six narrow regions of glass surrounding the core. The light chose to remain in these regions, despite the close proximity of large areas of silica, full of modes. One interpretation of this behavior is that, for particular wavelengths, the phase velocity of the light in the core is not coincident with any of the phase velocities available in the transmission band created by the array of larger adjacent regions. Thus light is unable to tunnel over to them and so remains trapped in the core. This "frustrated tunneling" mechanism of photonic band gap formation can only operate in extended two-dimensional structures (similar types of behavior are seen in modes trapped within defect layers in multilayer stacks[23]). The guided modes, which can also form in hollow-core PCF, can be viewed as surface-trapped states, light being confined by TIR on the core side of a boundary and by a photonic band gap on the cladding side. The zone of operation in Fig. 4 is $1 < n_z < n_{FSM}$.

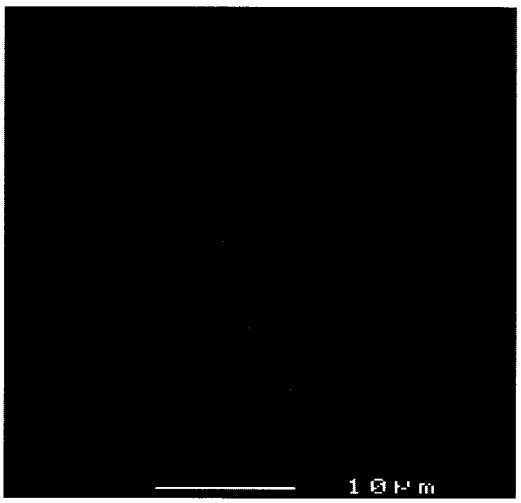


Fig. 5 Scanning electron micrograph of a hollow-core PCF designed to guide 1064 nm light. (Transmission spectrum is plotted in Fig. 8.) Note the slightly elliptical core.

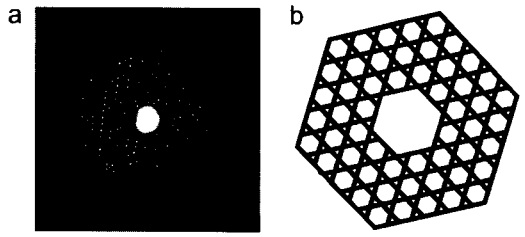


Fig. 6 (a) PCF designed for guidance of white light. (From Ref. [21].) (b) The "Kagomé" lattice in the cladding provides a low density of states (i.e., an incomplete band gap).

Low Density-of-States Guidance

It was shown that the transmission bands can be greatly widened by fabricating a different cladding structure—a Kagomé lattice.[21] Fig. 6a shows the white-light mode emerging from such a fiber, the minimum loss being ~1 dB/m. It turns out in this case that the density of states is greatly reduced in the cladding near the vacuum line. The consequential poor match between the core states and the greatly reduced number of cladding states slows down, but does not completely prevent, the leakage of light. The

precise details of the mechanism of this case are still under study.

CHARACTERISTICS OF GUIDANCE

Controlling Number of Modes

In a solid-core PCF guiding by modified TIR, the number of modes guided at short wavelengths is approximately given by:

$$N \approx \frac{V^2}{2} = \frac{2\pi^2\rho^2}{\lambda^2}(n_{co}^2 - n_{cl}^2) = \frac{z_{01}^2\rho^2}{2a^2} \quad (3)$$

where Eq. 1 has been used to represent the cladding refractive index. This indicates that in the short wavelength limit, the number is determined solely by geometry. This behavior can be understood by viewing the array of holes as a modal filter or "sieve" (Fig. 7). The fundamental mode has a transverse effective wavelength (in the core) $\lambda_{eff} \approx 4\Lambda$. Thus it is unable to "squeeze between" the gaps between the holes (in the cladding), which are spaced $\Lambda - d$ apart and are thus below the Rayleigh resolution limit $\lambda_{eff}/2 = 2\Lambda$. Provided the relative hole size d/Λ is small enough, higher-order modes are able to escape—their transverse effective wavelength is shorter so they have higher resolving power. As the holes are made larger, successive higher-order modes become trapped. Numerical modelling shows that if $d/\Lambda < 0.43$, the fiber is "endlessly single-mode" (i.e., it never supports any higher-order guided modes). The strong wavelength dispersion in the photonic crystal cladding causes the core-cladding index step to fall as the wavelength gets shorter.[17,18] This counteracts the usual trend toward increasingly multimode behavior at short wavelengths. In the limit of very short wavelength, the transverse single-mode profile freezes into a constant shape that does not depend on wavelength. In this regime, the light strikes the glass–air interfaces at glancing incidence, and is strongly rejected from the air holes. As a consequence, the angular divergence (numerical aperture or NA) of the emerging light is proportional to wavelength—in SMFs, it is approximately constant owing to the appearance of more and more higher-order guided modes as the frequency increases.

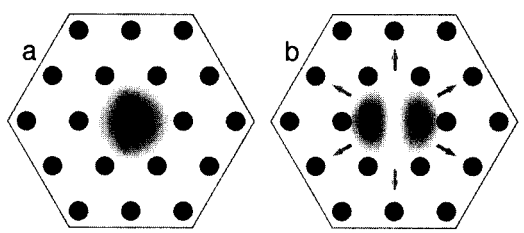


Fig. 7 Modal filtering in a solid-core PCF. (a) The fundamental mode is trapped whereas (b) higher-order modes leak away through the gaps between the air holes. (*View this art in color at www.dekker.com.*)

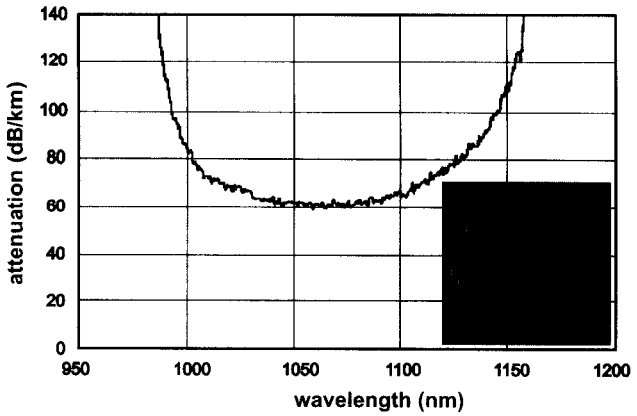


Fig. 8 Attenuation spectrum of a hollow-core PCF designed for low-loss transmission of 1064 nm light.

Ultralarge Mode Areas

Because this sieving process is controlled only by the size and positions of the holes and the core, it does not depend strongly on wavelength. A corollary is that the behavior is quite independent of the absolute size of the structure, making possible arbitrarily large SMF cores. A PCF with a core diameter of 22 μm at 458 nm was reported in 1998.[24] In conventional step index fibers, where $V < 2.405$ for single-mode operation, this would require uniformity of core refractive index to ~1 part in 10^5—this is very difficult if MCVD is used to form the doped core. Large-mode areas allow much higher power to be carried before the onset of intensity-related damage. This has obvious benefits for high-power delivery, amplifiers, and lasers.

Restricting the Guided Wavelengths

In solid-core PCF, the refractive index of the photonic crystal cladding increases with optical frequency, tending toward the index of silica glass in the short wavelength limit. Thus if the core is made from a glass with a refractive index lower than that of silica (e.g., fluorine-doped silica), guidance is lost at wavelengths shorter than a certain threshold value.[25] Such fibers have the unique ability to prevent transmission of short wavelength light—a remarkable result when one considers that conventional fibers get better and better at guiding light as the wavelength drops.

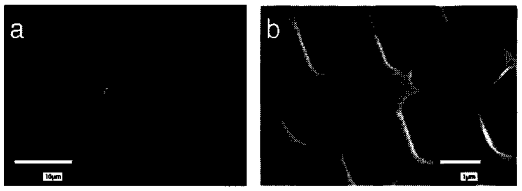


Fig. 9 (a) Scanning electron micrograph of birefringent PCF. (b) "Cobweb" PCF with very small core and zero GVD wavelength of 560 nm. The core is only 800 nm in diameter.

In hollow-core PCF, guidance can only occur when a photonic band gap coincides with a core resonance. This means that only restricted bands of wavelength are guided—colored light is typically transmitted when white light is launched into the core. The attenuation spectrum of a recent hollow-core PCF, designed for low-loss transmission at 1064 nm wavelength, is given in Fig. 8.[26]

Birefringence

The modes of a perfect sixfold symmetric core are not birefringent.[27] However, in practice, even slight distortions in the structure yield a degree of birefringence, and if the core is deliberately distorted so as to become twofold symmetric (e.g., by introducing capillaries with different wall thicknesses above and below the core; Fig. 9a), extremely high values of birefringence can be achieved—some 10 times larger than in conventional fibers.[28] Experiments show that the birefringence is highly insensitive to temperature, which is important in many applications. Traditional "polarization-maintaining" fibers (bow tie, elliptical core, or Panda) contain at least two different glasses, each with a different thermal expansion coefficient. The resulting temperature-dependent stresses make birefringence a strong function of temperature.

Dispersion

Group velocity dispersion (GVD), which causes different frequencies of light to travel at different speeds, is a crucial factor in the design of telecommunications systems. The magnitude of the GVD changes with wavelength, passing through zero at ~ 1.3 μm in conventional fiber. In PCF, the dispersion can be controlled with great freedom. For example, as the holes get larger, the core becomes more and more isolated, until it resembles an isolated strand of silica glass suspended by six thin webs of glass. If the whole structure is made very small (core diameters less than 1 μm have been made), the zero dispersion point can be shifted to wavelengths in the visible.[29,30] The "cobweb" PCF in Fig. 9b has a dispersion zero at 560 nm. By careful design, the wavelength dependence of the GVD can also be reduced. Fig. 10 shows the GVD curves of three fibers with ultraflattened low-level GVD.[31] In nonsilica glasses, such as Schott SF6, the intrinsic zero dispersion wavelength is at a difference point, providing a further degree of freedom in GVD design.[13]

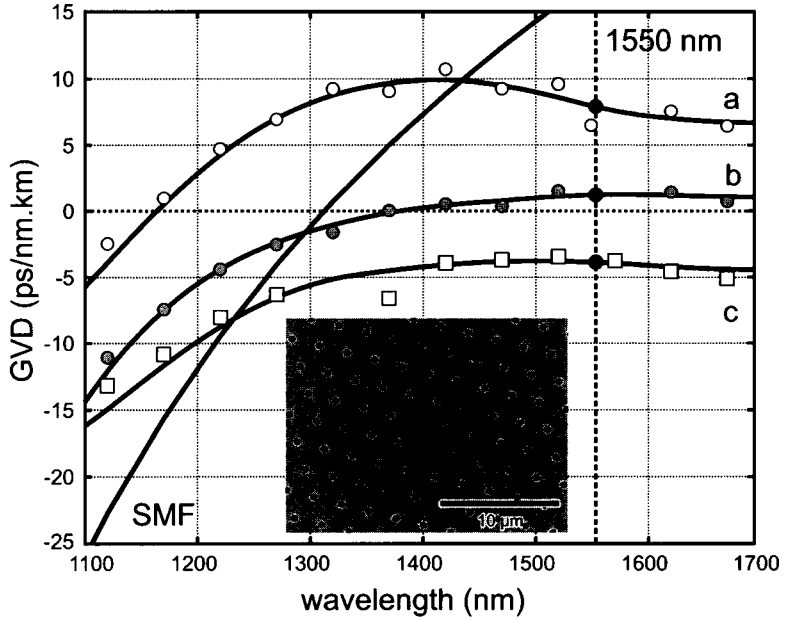


Fig. 10 GVD profiles, against wavelength, for three different PCFs designed to have low-level ultraflattened GVD. (From Refs. [31] and [32].)

Multicore

The stacking procedure makes it straightforward to produce multicore fibers. The preform stack is built up with the desired number of solid (or hollow) cores, and is drawn down to fiber in the usual manner.[33] The characteristics of coupling between the cores are unusual—the coupling strength depends on the sites chosen because the evanescent decay rate of the fields changes with azimuthal direction. Applications include curvature sensing.[34] More elaborate structures can be built up, such as fibers with inner and outer claddings designed as high-power fiber lasers.

Attenuation Mechanisms

The optimum spacing (~80 km) between repeaters in telecommunications systems is set by the fiber attenuation, which in the best conventional fibers is limited by Rayleigh scattering to ~0.2 dB/km at 1550 nm wavelength. A key question is whether hollow-core or solid-core PCF can match—or perhaps improve on—this and perhaps replace rival conventional fibers in telecommunications. The answer is not yet clear, although it is easy enough to pose the questions. Are the glass–air interfaces smooth enough to avoid significant scattering out of the core or into other modes? Is Rayleigh scattering amplified by the large refractive index step at the interfaces? Will the holes fill with water vapor and will huge water-related losses develop at 1.39 μm wavelength, where an overtone of the OH bond absorption occurs?

Absorption and scattering

The reported losses, from teams in Europe, Japan, and the United States, are steadily dropping, and the record now stands at 0.37 dB/km in a solid-core PCF.[35,36] Hollow-core PCF has the greatest potential for extremely low loss because the light is travelling predominantly in the hollow core. Although the best reported attenuation in hollow-core PCF is fairly high (13 dB/km),[37] values well below 0.2 dB/km seem at least feasible with further development of the technology. The prospect of improving on conventional fiber, at the same time greatly reducing the nonlinearities associated with a solid glass core, is tantalizing.

Bend loss

Conventional fibers suffer additional loss if bent beyond a certain critical radius R_{crit}, which depends on wavelength, core-cladding refractive index step, and, most notably, the third power of core radius a^3.[38] For wavelengths longer than a certain value (the "long wavelength bend edge"), all guidance is effectively lost. PCF does not escape this effect (large-mode area PCF in particular is highly sensitive to bend loss) and, in fact, in its endlessly single-mode form, PCF exhibits an unexpected *short* wavelength bend edge, caused by bend-induced coupling from fundamental to higher-order modes, which of course leak out of the core.[17] Experimentally, hollow-core PCF is remarkably insensitive to bend loss—in many cases, no appreciable drop in transmission is observed until the fiber breaks.

Confinement loss

A fabricated PCF necessarily incorporates a finite number of holes in the cladding region. For a guided mode, the Bloch waves in the photonic crystal cladding are evanescent, just like the evanescent plane waves in the cladding of a conventional fiber. If the cladding is not thick enough, the evanescent fields at the cladding/coating boundary can be substantial, causing attenuation. In the solid-core case for small values of d/Λ, the resulting loss can be quite substantial unless a large number of periods is used.[31] Very similar losses are observed in the hollow-core fibers, where the "strength" of the photonic band gap (closely related to its width in β) determines how many periods are needed to reduce confinement loss to acceptable levels. Numerical modelling is very useful for giving an indication of how many periods are needed to reach a required loss level.

NUMERICAL MODELING TECHNIQUES

The complex structure of PCF—in particular the large refractive index difference between glass and air and the resonant nature of the holes and interhole spaces—makes its electromagnetic analysis very challenging. Standard optical fiber analyses do not much help. Although—as we have seen in previous sections—a number of approximate models do exist, these are only useful as rough guidelines to the exact behavior. Maxwell's equations must be solved numerically, using one of a number of specially developed techniques.[16,39–42]

The most popular technique uses a plane wave basis for the fields. This converts the wave equations into standard matrix eigenvalue problems suitable for numerical computation. The wave equation satisfied by the **H** field is particularly amenable because, set up for (ω^2/c^2) as an eigenvalue, it is Hermitian in form. For waveguiding structures, such as PCFs that incorporate a defect core into an otherwise periodic medium, a supercell is constructed. The supercell is made large enough so that, once tiled, the guided modes in different cores do not significantly interact. The plane wave expansion method suffers from its

inability to accurately represent step changes in the dielectric function. The associated ringing or Gibb's phenomenon introduces an inaccuracy in the solutions unless a very large number of plane wave components is retained. This problem can be alleviated by reexpressing the wave equation satisfied by the **H** field in a variational form and by using a conjugate gradient minimization scheme.[43]

In most practical cases in photonics, a set of constant frequency modes is far more useful. Therefore it is convenient to set up Maxwell's equation with β^2 as an eigenvalue:

$$\left(\nabla_{\|}^2 + \frac{\omega^2}{c^2}\varepsilon(\mathbf{r}_{\|})\right)\mathbf{H}_{\|} + [\nabla_{\|} \ln \varepsilon(\mathbf{r}_{\|})] \wedge (\nabla_{\|} \wedge \mathbf{H}_{\|}) = \beta^2 \mathbf{H}_{\|} \quad (4)$$

where the gradient operator within the square brackets only operates on the $\ln \varepsilon(\mathbf{r}_{\|})$ term. This form also allows material dispersion to be naturally included. Fourier expansion of Eq. 4 can then be used to convert them to matrix eigenvalue forms.[40] Unfortunately, because of the non-Hermitian nature of Eq. 4, a variational approach cannot be invoked so straightforwardly, implying that the number of retained basis waves that can be considered in a plane wave expansion is more limited.

Other numerical techniques include expansion of the field in terms of Hermite–Gaussian functions,[29] the finite-difference time domain (FDTD) approach (a simple and versatile tool with which to explore waveguide geometries[44]), and the finite-element approach.[45] If the PCF structure comprises purely circular holes, the multipole or Rayleigh method is a particularly fast and efficient approach to use.[42] The method works by calculating the scattering of a field incident on each hole by the use of Mie scattering theory.

The most appropriate and efficient method to use for a given PCF structure depends on its geometry and which properties are required. The plane wave method is attractive for calculating mode properties because of the large number of basis waves that can be used. In common with most basis expansion methods, this method fails to accurately describe the discontinuities in the dielectric distribution and in the field at the dielectric interfaces; this necessitates the use of a large basis size. The multipole method overcomes this limitation by expanding the fields in various dielectric regions in local basis sets appropriate to the local symmetry and performs the boundary matching directly. For fibers comprising circular holes, the multipole method generally gives the quickest and most accurate results. Because no spatial mesh is required, the method naturally maintains the symmetry of the structure and the constructed eigenmodes show the correct symmetries also.

The confinement loss suffered by the guided modes is not directly accessible with the plane wave approach because of the Bloch boundary conditions that are applied. The leakage can be obtained from methods in which radiative boundary conditions outside the holey region can accurately be employed. This is automatic for the multipole method and can be achieved using boundary-element and finite-element approaches, although in the latter case, its accurate implementation (using, for example, perfectly matched layers) is not straightforward. The finite difference time domain approach can be invoked to calculate leakage loss if this is substantial, but losses at the level of a few decibels per kilometer or less, which are appropriate for most PCF fiber systems, are not resolved. This problem can be alleviated by reexpressing the wave equation satisfied by the **H** field in a variational form and by using the conjugate gradient minimization scheme;[43] this allows many tens of thousands of plane waves to be retained in the expansions.

INTRAFIBER DEVICES, CUTTING, AND JOINING

As PCF becomes more widely used, there is increasing demand for cleaving, low-loss splicing, multiport couplers, intrafiber devices, and mode area transformers. The air holes provide an opportunity not available in standard fibers—they collapse under surface tension when heated to the softening temperature of the glass. Thus not only can the fiber be stretched locally to reduce its cross-sectional area, but the microstructure can itself be altered.

PCF cleaves very cleanly using standard tools, showing slight end-face distortion only when the core crystal is extremely small (interhole spacing ~ 1 μm), and the air-filling fraction is very high (>50%). Solid-core PCF can be spliced successfully both to itself and to standard fiber using a fusion-splicing process. The two fiber ends are placed in intimate contact and heated to softening point. With careful control, they fuse together without distortion. Provided that the mode areas are well matched, splice losses of <0.2 dB can normally be achieved except when the core is extremely small (less than ~ 1.5 μm). Fusion splicing hollow-core fiber is feasible when there is a thick solid glass outer sheath (e.g., as depicted in the inset in Fig. 8), although very low splice losses can be obtained simply by placing the fibers end-to-end and clamping them—after all, the index matching fluid for hollow-core PCF is vacuum.

In many applications, it is important to be able to change the mode area without losing light. This is done traditionally using miniature bulk optics—tiny lenses precisely designed to match to a desired NA and spot size. In PCF, an equivalent effect can be obtained by scanning a

heat source (flame or carbon dioxide laser) along the fiber. This causes the holes to collapse, with the degree of collapse depending on the dwell time of the heat. Drawing the two fiber ends apart at the same time provides additional control. Low loss graded transitions can fairly easily be made—mode diameter reductions as high as 5:1 have been realized with low loss. A further degree of freedom may be gained by pressurizing the holes during the taper process.[46]

The ability to cause large changes in the optical characteristics of PCF by heating is leading to a whole family of new intrafiber components. Microcouplers can be made in dual-core PCF—two optically isolated cores can be made to interact by collapsing the holes so as to allow the mode fields to expand and "talk" to each other.[47] Long period gratings, which couple the core light into the cladding within certain wavelength bands, can be made by periodic modulation of hole size.[48] Rocking a birefringent PCF to-and-fro, while scanning a carbon dioxide laser along it, results in so-called "rocking filters," which transfer power from one polarization state to the other within a narrow band of wavelengths.[49] All these components have one great advantage over equivalent devices made in conventional fiber—being permanent changes in morphology, they are highly stable with temperature and over time.

APPLICATIONS

The diversity of new or improved features, beyond what conventional fiber offers, means that PCFs are finding an increasing number of applications in every widening area of science and technology.

Gas-Based Nonlinear Optics

A long-standing challenge in photonics is how to maximize nonlinear interactions between laser light and low-density media such as gases. Efficient nonlinear processes require high intensities at low power, long interaction lengths, and good-quality transverse beam profiles. A structure conceptually capable of delivering all these requirements simultaneously is a perfectly guiding hollow-core waveguide supporting a single transverse mode with low attenuation losses. Although, theoretically, this could be realized using a perfect metal, the attenuation in real metals at optical frequencies is much too high. A number of conventional approaches have been used to alleviate this problem, including focusing a laser beam into the gas with suitable optics, using a $\sim$200-μm bore fiber capillary to confine the gas and provide some degree of guidance for the light,[50] and employing a gas-filled high-finesse Fabry–Pérot cavity to increase the interaction length.[51] As explained in Ref. [21], none of these approaches comes close to the performance offered by hollow-core PCF. At a bore diameter of 10 μm, for example, a focused free space laser beam is marginally preferable to a capillary, whereas a hollow-core PCF with 13 dB/km attenuation is an astonishing 100,000$\times$ more effective. Such huge enhancements are rare in physics, and point the way to dramatic improvements in all sorts of nonlinear laser–gas interactions. The first results are beginning to emerge: Recently stimulated Raman scattering was observed in a hydrogen-filled hollow-core PCF at threshold pulse energies $\sim$100$\times$ lower than previously reported.[21] Many other areas of nonlinear optics are likely to be revolutionized, such as x-ray generation in noble gases pumped by femtosecond Ti–sapphire laser pulses.[52] The conversion efficiency of this process was further enhanced recently by modulating the bore diameter so as to phase match the light and the x-rays.[53] This can be implemented in PCF by heat treatment with carbon dioxide laser light.[48]

High-Power Transmission

Hollow-core fiber is also excellent for transmitting high continuous-wave power as well as ultrashort pulses with very high peak powers. A recent paper reports femtosecond solitons at 1550 nm with peak powers of 2 MW in hollow-core PCF.[54] For soliton formation, the sign and magnitude of the GVD are crucial. It has recently been reported that the GVD changes sign across the band gap, permitting choice of normal or anomalous dispersion depending on the application.[22]

Rare Earth-Doped Lasers and Amplifiers

Rare earth-doped PCF lasers can be straightforwardly produced by incorporating doped canes in the preform stack. Many different structures can be fashioned, such as large-mode area cores and multicore lasers.

Ultrahigh power lasers

High-power rare earth-doped lasers and amplifiers can be fashioned in large-mode area PCF by incorporating a second, much larger, multimode core around the single-mode lasing core. Using microstructuring techniques, this "inner cladding waveguide" can be suspended by connecting it to an outer glass tube with very thin webs of glass (Fig. 11).[56] Very high inner cladding NA can be obtained (>0.9), making it easy to launch and guide light from high-power diode bar pump lasers, which typically have very poor beam quality. The pump light is absorbed by the lasing core, and high-power single-mode operation

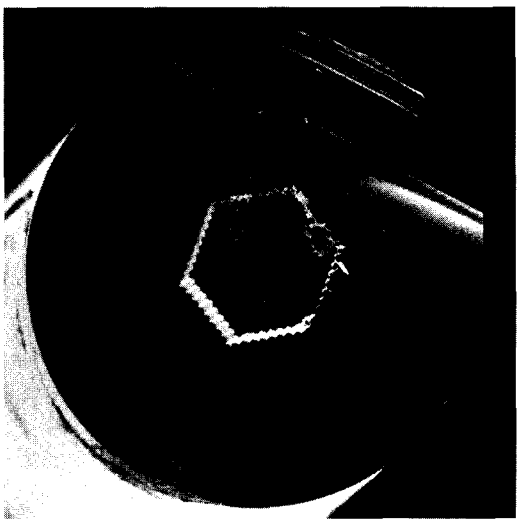


Fig. 11 Scanning electron micrograph of a cladding-pumped PCF laser designed for high-power operation. (From Ref. [55].) The lasing core is a missing air hole close to the left-hand corner of the photonic crystal.

can be achieved, as in recent reports for Yb-doped fiber lasers.[55,57]

Mode-locked fiber lasers

Mode-locked fiber lasers rely on control of the dispersion within the laser cavity to achieve ultrashort pulse (subpicosecond) operation. PCF offers fine-tuning of the GVD over a wide wavelength range, making it ideal for developing a new generation of mode-locked laser systems. Laser cavities can be envisaged in which there is high local GVD but very low average dispersion over a broad wavelength range. This would allow operation in the stretched-pulse regime, or could be used simply to cancel the dispersion of other cavity components. Mode-locked fiber lasers previously were forced to operate at wavelengths beyond the 1.3-μm zero GVD point, where the dispersion in SMF is anomalous. However, using PCF, it becomes possible to design all-fiber mode-locked systems at shorter wavelengths by shifting the zero GVD wavelength down into the visible.

Atom and Particle Guidance

A focused light beam produces both a longitudinal (accelerating) and a transverse (trapping) force on dielectric microparticles.[58] For maximum trapping force, the intensity gradient of the light must be as high as possible. This can be achieved by focusing a high-power laser beam with a high NA lens. The Rayleigh length for such a tightly focused beam is rather short; furthermore, the high intensity will quickly accelerate the particle out of the trapping zone. Stable trapping of a particle in free space requires the longitudinal force to be balanced either by gravity or by a second (or reflected) beam. It has long been recognized that an alternative configuration offering stability would be a nondiffracting or guided beam. To guide a dielectric particle, the beam would have to be trapped in air rather than in glass, which, until very recently, was not possible. That did not stop the demonstration of guidance of both solid particles[59] and atoms[60] along a length of hollow capillary during the past decade. Of course, use of a capillary prevented the full exploitation of the possibilities of a hollow waveguide—for the small core sizes needed for strong transverse confinement, capillary losses are very high, even on length scales of a few millimeters. The same is not true in

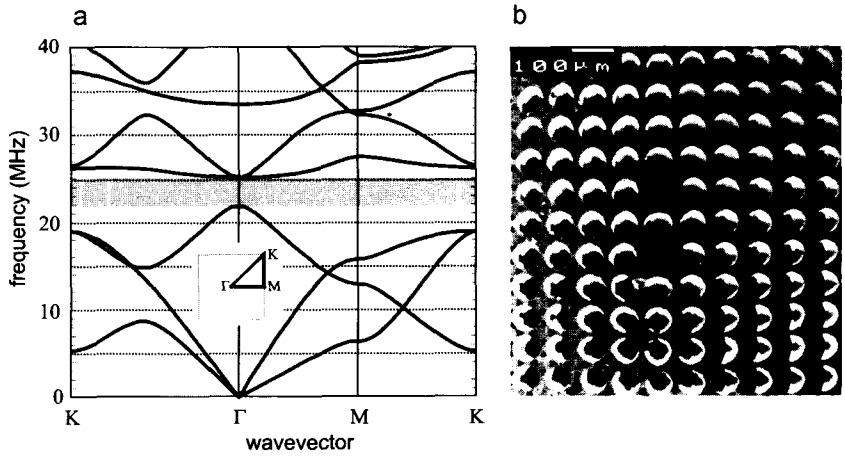


Fig. 12 (a) Photonic band gaps of a square lattice PCF preform. (b) Full two-dimensional acoustic band gaps occur around 23 MHz, in agreement with the experiments. (From Ref. [64].)

hollow-core PCF, where core diameters of ~ 10 μm can be realized with losses as low as 50 dB/km. Such a small spot size gives a strong transverse trapping force for a given longitudinal force, so that it becomes possible to envisage guiding particles even around tightly bent fibers at reasonably high speeds. The first steps toward achieving this goal were recently taken.[61]

Sound–Light Interactions

A perennial problem in photonics is how to force sound and light to "talk" (i.e., interact more strongly). Although in theory a single photon can be deflected by annihilation or creation of a single phonon, a fair population of phonons must be used to swamp out the effects of thermal vibrations. Even so, commercial acousto-optic Bragg cells are fantastically inefficient. Driven at 500 mW power and deflecting a laser beam of 1000 nm wavelength and a power of 1 mW, a bulk 40-MHz Bragg cell uses roughly 4 billion phonons to deflect one photon. This huge overkill is largely because of the very poor overlap between sound and light. One way of improving this is to microstructure the material so as to create a frequency band where it falls silent—an acoustic or photonic band gap.[62,63] A suitable defect volume, placed within this material, will resonate at frequencies within the band gap, with the acoustic waves being unable to escape into the surrounding silence. Very high acoustic energy densities can then be reached at small drive powers, suggesting the possibility of a new class of superefficient acousto-optic component in which both sound and light are trapped at the same time in the same place. Such components would enjoy radically enhanced performance (in terms of drive power and precision). The first steps in this direction were taken recently with the demonstration of intraphotonic band gap trapped resonances in the core of a square lattice PCF preform[64] (Fig. 12).

Nonlinear Effects

The ability to enhance effective nonlinearity and, at the same time, control the magnitude and wavelength dependence of the GVD makes PCF a highly versatile and complete "laboratory-in-a-fiber" for nonlinear optics. Many well-known nonlinear fiber optical effects (e.g., stimulated Raman scattering, four-wave mixing, self-phase modulation, modulation instability, and solitonic effects) have already been observed in short (few meters) lengths of fiber. The parameter:

$$\gamma = \omega_0 n_2 / c A_{\text{eff}} \tag{5}$$

is often used as a measure of the nonlinearity of the fiber.[65] In Eq. 5, n_2 is the nonlinear refractive index of the fiber material (2.5×10^{-20} m^2 W^{-1} in silica glass, and an order of magnitude or more higher for soft glasses) and A_{eff} is the effective area of the guided mode. The highest nonlinearity available in conventional silica fibers is $\gamma \sim 20$ W^{-1} km^{-1} at 1550 nm wavelength[66] and $\gamma \sim 10$ W^{-1} km^{-1} at 850 nm. In comparison, $\gamma \sim 240$ W^{-1} km^{-1} at 850 nm for the PCF in Fig. 13 and nonlinearities as high as 550 W^{-1} km^{-1} at 1550 nm have been measured for PCFs based on soft glasses.[67] The fiber nonlinearity is important because it sets the nonlinear phase shift associated with propagation of light of a given intensity. The nonlinearity can be used to calculate the nonlinear length of the fiber $L_{\text{nl}} = (\gamma P_0)^{-1}$, where P_0 is the peak power. For example, for the fiber in Fig. 13, a peak power of 10 kW (easily obtained from an unamplified ultrashort pulse laser system) implies $L_{\text{nl}} < 0.5$ mm. For typical values of loss (say, 100 dB/km) and dispersion (e.g., 10 psec/nm km), this nonlinear length is far shorter than the fiber effective length L_{eff} and the dispersive length L_{D}.[65] Consequently, it is easy to observe extraordinary nonlinear effects in such fibers. The actual nonlinear processes that dominate in a particular experiment are determined by the GVD as well as the laser characteristics.

Perhaps the most dramatic application of nonlinear PCF is to supercontinuum (SC) generation from picosecond and femtosecond laser pulses. When high-power pulses travel through a material their frequency spectrum can be broadened by a range of interconnected nonlinear effects.[68] Until recently, the preferred pump laser was a regeneratively amplified Ti–sapphire system producing high (millijoules) energy femtosecond pulses at 800 nm wavelength and kilohertz repetition rate. Then in 2000, it was observed that highly nonlinear PCF, designed with zero GVD close to 800 nm, massively broadens the

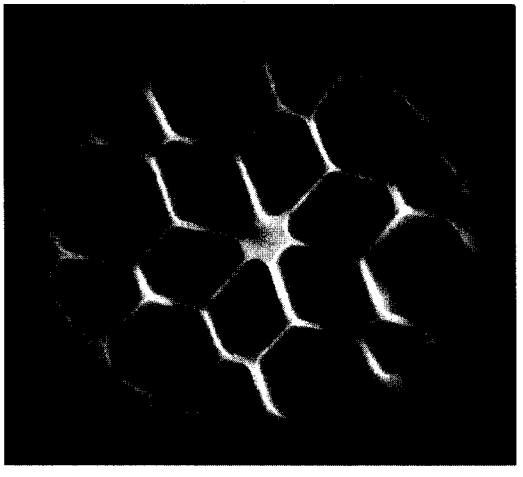


Fig. 13 Core region of a silica/air fiber with a 1-μm core diameter. Its unusual nonlinear and dispersive properties make it a spectacular vehicle for studying nonlinear fiber optics.

spectrum of low-energy (few nanojoules) unamplified Ti–sapphire pulses launched into just a few centimeters of fiber.[69,70] Removal of the need for a power amplifier, the hugely increased 100-MHz repetition rate, and the spatial and temporal coherence of the light emerging from the core explain why this source is unique. The SC has applications in optical coherence tomography,[71] frequency metrology,[72] and all kinds of spectroscopy.

SC has been generated in different PCFs at many pump wavelengths in the range of 648 nm[73] to 1550 nm.[13] The wide range of GVD characteristics that have been demonstrated[31] in PCF in the 1550-nm wavelength band has provided a test bed[32] for nonlinear fiber optics, in which the sensitivity of nonlinear pulse propagation to the details of the dispersion curve can be tested. The development of conventional fiber-based optical parametric oscillators and amplifiers has been constrained by the very limited scope for engineering the GVD profile. In PCF, these constraints are lifted. The small effective mode area also offers high gain, and PCF-based oscillators based on synchronously pumped femtosecond and picosecond pump pulses have been demonstrated with relatively low powers.[74–76]

CONCLUSION

PCFs represent a next-generation, radically improved version of a well-established and highly successful technology. In escaping from the confines of conventional fiber optics, they have created a renaissance of new possibilities in many diverse areas of research and technology, in the process irrevocably breaking many of the tenets of received fiber optics wisdom. Over the next few years, we expect to see a gradual takeup of PCFs in many fields of application. Emerging fields of application include biotechnology, optical sensing, medical physics (imaging techniques, surgery), microparticle delivery, high-power fiber lasers, fiber-based laser delivery for manufacturing, long-haul telecommunications, and gas-based nonlinear wavelength conversion.

REFERENCES

1. Colladon, D. Comptes Rendus October 24, **1842**, *15*, 800–802.
2. Nagel, S.R.; MacChesney, J.B.; Walker, K.L. An overview of the modified chemical vapor deposition (MCVD) process and performance. IEEE J. Quantum Electron. **1982**, *18*, 459–476.
3. Yeh, P.; Yariv, A. Bragg reflection waveguides. Opt. Commun. **1976**, *19*, 427–430.
4. Brechet, F.; Roy, P.; Marcou, J.; Pagnoux, D. Single mode propagation into depressed-core-index photonic-bandgap fibre designed for zero-dispersion propagation at short wavelengths. Electron. Lett. **2000**, *36*, 514–515.
5. Johnson, S.G.; Ibanescu, M.; Skorobogatiy, M.; Weisberg, O.; Engeness, T.D.; Soljacic, M.; Jacobs, S.A.; Joannopoulos, J.D.; Fink, Y. Low-loss asymptotically single-mode propagation in large-core OmniGuide fibres. Opt. Express **2001**, *9*, 748–779.
6. Russell, P.St.J. **1991**. private papers.
7. Knight, J.C.; Birks, T.A.; Russell, P.St.J. Holey Silica Fibres. In *Optics of Nanostructured Materials*; Markel, V.A., George, T.F., Eds.; John Wiley and Sons: New York, 2001; 39–71.
8. Birks, T.A.; Knight, J.C.; Mangan, B.J.; Russell, P.St.J. Photonic crystal fibres: An endless variety. IEICE Trans. Electron. **2001**, *E84-C*, 585–592.
9. Russell, P.St.J. Photonic crystal fibers. Science **2003**, *299*, 358–362.
10. Knight, J.C.; Birks, T.A.; Russell, P.St.J.; Atkin, D.M. In *Pure Silica Single-Mode Fibre with Hexagonal Photonic Crystal Cladding*, Optical Fibre Communications Conference (OFC'96), San Diego, March 1996. postdeadline paper PD3.
11. Knight, J.C.; Birks, T.A.; Russell, P.St.J.; Atkin, D.M. All-silica single-mode fibre with photonic crystal cladding. Opt. Lett. **1996**. *21*, 1547–1549. Errata; Opt. Lett., **1997**, *22*, 484–485.
12. Allan, D.C.; West, J.A.; Fajardo, J.C.; Gallagher, M.T.; Koch, K.W.; Borrelli, N.F. Photonic Crystal Fibres: Effective Index and Bandgap Guidance. In *Photonic Crystals and Light Localisation in the 21st Century*; Soukoulis, C.M., Ed.; Kluwer Academic Publishers, 2001; 305–320.
13. Ravi Kanth Kumar, V.V.; George, A.K.; Reeves, W.H.; Knight, J.C.; Russell, P.St.J.; Omenetto, F.G.; Taylor, A.J. Extruded soft glass photonic crystal fiber for ultrabroad supercontinuum generation. Opt. Express **2002**, *10*, 1520–1525.
14. Kiang, K.M.; Frampton, K.; Monro, T.M.; Moore, R.; Tucknott, J.; Hewak, D.W.; Richardson, D.J.; Rutt, H.N. Extruded single-mode non-silica glass holey optical fibres. Electron. Lett. **2002**, *38*, 546–547.
15. van Eijkelenborg, M.A.; Large, M.C.J.; Argyros, A.; Zagari, J.; Manos, S.; Issa, N.A.; Bassett, I.; Fleming, S.; McPhedran, R.C.; de Sterke, C.M.; Nicorovici, N.A.P. Microstructured polymer optical fibre. Opt. Express **1999**, *9*, 319–327.
16. Birks, T.A.; Roberts, P.J.; Russell, P.St.J.; Atkin,

D.M.; Shepherd, T.J. Full 2-D photonic band gaps in silica/air structures. Electron. Lett. **1995**, *31*, 1941–1942.
17. Birks, T.A.; Knight, J.C.; Russell, P.St.J. Endlessly single-mode photonic crystal fibre. Opt. Lett. **1997**, *22*, 961–963.
18. Knight, J.C.; Birks, T.A.; Russell, P.St.J.; de Sandro, J.-P. Properties of photonic crystal fibre and the effective index model. J. Opt. Soc. Am., A **1998**, *15*, 748–752.
19. Cregan, R.F.; Mangan, B.J.; Knight, J.C.; Birks, T.A.; Russell, P.St.J.; Roberts, P.J.; Allan, D.C. Single-mode photonic band gap guidance of light in air. Science **1999**, *285*, 1537–1539.
20. Knight, J.C.; Broeng, J.; Birks, T.A.; Russell, P.St.J. Photonic band gap guidance in optical fibres. Science **1998**, *282*, 1476–1478.
21. Benabid, F.; Antonopoulos, G.; Knight, J.C.; Russell, P.St.J. Stimulated Raman scattering in hydrogen-filled hollow-core photonic crystal fiber. Science **2002**, *298*, 399–402.
22. Bouwmans, G.; Luan, F.; Knight, J.C.; Russell, P.St.J.; Farr, L.; Mangan, B.J.; Sabert, H. Properties of a hollow-core photonic bandgap fiber at 850 nm wavelength. Opt. Express **2003**, *11*, 1613–1620.
23. Pechstedt, R.D.; Russell, P.St.J. Narrow-band in-line fibre filter using surface-guided Bloch modes supported by dielectric multilayer stacks. J. Lightwave Technol. **1996**, *14*, 1541–1545.
24. Knight, J.C.; Birks, T.A.; Cregan, R.F.; Russell, P.St.J.; de Sandro, J.-P. Large mode area photonic crystal fibre. Electron. Lett. **1998**, *34*, 1347–1348.
25. Mangan, B.J.; Arriaga, J.; Birks, T.A.; Knight, J.C.; Russell, P.St.J. Fundamental-mode cutoff in a photonic crystal fibre with a depressed-index core. Opt. Lett. **2001**, *26*, 1469–1471.
26. www.blazephotonics.com.
27. Steel, M.J.; White, T.P.; de Sterke, C.N.; McPhedran, R.C.; Botten, L.C. Symmetry and degeneracy in microstructured optical fibers. Opt. Lett. **2001**, *26*, 488–490.
28. Ortigosa-Blanch, A.; Knight, J.C.; Wadsworth, W.J.; Arriaga, J.; Mangan, B.J.; Birks, T.A.; Russell, P.St.J. Highly birefringent photonic crystal fibres. Opt. Lett. **2000**, *25*, 1325–1327.
29. Mogilevtsev, D.; Birks, T.A.; Russell, P.St.J. Group-velocity dispersion in photonic crystal fibres. Opt. Lett. **1998**, *23*, 1662–1664.
30. Knight, J.C.; Arriaga, J.; Birks, T.A.; Ortigosa-Blanch, A.; Wadsworth, W.J.; Russell, P.St.J. Anomalous dispersion in photonic crystal fibres. IEEE Photonics Technol. Lett. **2000**, *12*, 807–809.
31. Reeves, W.H.; Knight, J.C.; Russell, P.St.J.; Roberts, P.J. Demonstration of ultra-flattened dispersion in photonic crystal fibres. Opt. Express **2002**, *10*, 609–613.
32. Reeves, W.H.; Skryabin, D.V.; Biancalana, F.; Knight, J.C.; Russell, P.St.J.; Omenetto, F.G.; Efimov, A.; Taylor, A.J. Transformation and control of ultra-short pulses in dispersion-engineered photonic crystal fibres. Nature **2003**, *424*, 511–515.
33. Mangan, B.J.; Knight, J.C.; Birks, T.A.; Russell, P.St.J.; Greenaway, A.H. Experimental study of dual-core photonic crystal fibre. Electron. Lett. **2000**, *36*, 1358–1359.
34. MacPherson, W.N.; Gander, M.J.; McBride, R.; Jones, J.D.C.; Blanchard, P.M.; Burnett, J.G.; Greenaway, A.H.; Mangan, B.J.; Birks, T.A.; Knight, J.C.; Russell, P.St.J. Remotely addressed optical fibre curvature sensor using multicore photonic crystal fibre. Opt. Commun. **2001**, *193*, 97–104.
35. Farr, L.; Knight, J.C.; Mangan, B.J.; Roberts, P.J. In *Low Loss Photonic Crystal Fibre*, European Conference on Optical Communications ECOC'02, Copenhagen, 2002. postdeadline paper PD1.3.
36. Tajima, K.; Zhou, J.; Nakajima, K.; Sato, K. In *Ultra Low Loss and Long Length Photonic Crystal Fiber*, Optical Fiber Communication Conference (OFC 2003), Atlanta, March 2003. postdeadline paper PD1.
37. Smith, C.M.; Venkataraman, N.; Gallagher, M.T.; Mueller, D.; West, J.A.; Borrelli, N.F.; Allan, D.C.; Koch, K.W. Low-loss hollow-core silica/air photonic bandgap fibre. Nature **2003**, *424*, 657–659.
38. Snyder, A.W.; Love, J.D. *Optical Waveguide Theory*; Chapman and Hall: London, 1983.
39. Mogilevtsev, D.; Birks, T.A.; Russell, P.St.J. Localised function method for modelling defect modes in 2-D photonic crystals. IEEE J. Lightwave Technol. **1999**, *17*, 2078–2081.
40. Ferrando, A.; Miret, J.J.; Silvestre, E.; Andrés, P.; Andrés, M.V. Full-vector analysis of a realistic photonic crystal fibre. Opt. Lett. **1999**, *24*, 276–278.
41. Roberts, P.J.; Shepherd, T.J. The guidance properties of multi-core photonic crystal fibres. J. Opt. A, Pure Appl. Opt. **2001**, *3*, S1–S8.
42. McPhedran, R.C.; Botten, L.C.; Asatryan, A.A.; Nicorovici, N.A.; Robinson, P.A.; de Sterke, C.M. Calculation of electromagnetic properties of regular and random arrays of metallic and dielectric cylinders. Phys. Rev., E **1999**, *60*, 7614–7617.
43. Meade, R.D.; Rappe, A.M.; Brommer, K.D.; Joannopoulos, J.D.; Alerhand, O.L. Accurate theoretical

analysis of photonic band-gap materials. Phys. Rev., B **1993**, *48*, 8434–8437.
44. Chan, C.T.; Yu, Q.L.; Ho, K.M. Order *N* spectral method for electromagnetic waves. Phys. Rev., B **1995**, *51*, 16635–16642.
45. Mias, C.; Webb, J.P.; Ferrari, R.L. Finite element modelling of electromagnetic waves in doubly and triply periodic structures. IEE Proc. Optoelectron. **1999**, *146*, 111–118.
46. Birks, T.A.; Kakarantzas, G.; Russell, P.St.J.; Murphy, D.F. Photonic crystal fibre devices. Proc. SPIE **2002**, *4943*, 142–151. (Fiber-Based Component Fabrication, Testing and Connectorization).
47. Kakarantzas, G.; Dimmick, T.E.; Birks, T.A.; Le Roux, R.; Russell, P.St.J. Miniature all-fiber devices based on CO_2 laser micro-structuring of tapered fibers. Opt. Lett. **2001**, *26*, 1137–1139.
48. Kakarantzas, G.; Mangan, B.J.; Birks, T.A.; Knight, J.C.; Russell, P.St.J. In *Directional Coupling in a Twin Core Photonic Crystal Fiber Using Heat Treatment*, Conference on Lasers and Electro-Optics (CLEO), Baltimore, MD, May 2001. paper JTuD2.
49. Kakarantzas, G.; Ortigosa-Blanch, A.; Birks, T.A.; Russell, P.St.J.; Farr, L.; Couny, F.; Mangan, B.J. Structural rocking filters in highly birefringent photonic crystal fiber. Opt. Lett. **2003**, *28*, 158–160.
50. Rabinowitz, P.; Stein, A.; Brickman, R.; Kaldor, A. Efficient tunable hydrogen Raman laser. Appl. Phys. Lett. **1979**, *35*, 739–741.
51. Meng, L.S.; Repasky, K.S.; Roos, P.A.; Carlsten, J.L. Widely tunable continuous-wave Raman laser in diatomic hydrogen pumped by an external-cavity diode laser. Opt. Lett. **2000**, *25*, 472–474.
52. Brabec, T.; Krausz, F. Intense few-cycle laser fields: Frontiers of nonlinear optics. Rev. Mod. Phys. **2000**, *72*, 545–591.
53. Paul, A.; Bartels, R.A.; Tobey, R.; Green, H.; Weiman, S.; Christov, I.P.; Murnane, M.M.; Kapteyn, H.C.; Backus, S. Quasi-phase-matched generation of coherent extreme-ultraviolet light. Nature **2003**, *421*, 51–54.
54. Ouzounov, D.G.; Ahmad, F.R.; Gaeta, A.L.; Müller, D.; Venkataraman, N.; Gallagher, M.T.; Koch, K.W. In *Dispersion and Nonlinear Propagation in Air-Core Photonic Band-Gap Fibers*, Conference on Lasers and Electro-Optics (CLEO'03), Baltimore, June 2003. paper CThT5.
55. Wadsworth, W.J.; Percival, R.M.; Bouwmans, G.; Knight, J.C.; Russell, P.St.J. High power air-clad photonic crystal fibre laser. Opt. Express **2003**, *11*, 48–53.
56. Wadsworth, W.J.; Percival, R.M.; Bouwmans, G.; Knight, J.C.; Birks, T.A.; Hedley, T.D.; Russell, P.St.J. Very high numerical aperture fibres. IEEE Phot. Tech. Lett., March **2004**, *in press*.
57. Limpert, J.; Schreiber, T.; Nolte, S.; Zellmer, H.; Tünnermann, A.; Iliew, R.; Lederer, F.; Broeng, J.; Vienne, G.; Petersson, A.; Jakobsen, C. High-power air-clad large-mode-area photonic crystal fiber laser. Opt. Express **2003**, *11*, 818–823.
58. Ashkin, A. Acceleration and trapping of particles by radiation pressure. Phys. Rev. Lett. **1970**, *24*, 156–159.
59. Renn, M.J.; Pastel, R.; Lewandowski, H.J. Laser guidance and trapping of mesoscale particles in hollow-core optical fibres. Phys. Rev. Lett. **1999**, *82*, 1574–1577.
60. Renn, M.J.; Montgomery, D.; Vdovin, O.; Anderson, D.Z.; Wieman, C.E.; Cornell, E.A. Laser-guided atoms in hollow-core optical fiber. Phys. Rev. Lett. **1995**, *75*, 3253–3256.
61. Benabid, F.; Knight, J.C.; Russell, P.St.J. Particle levitation and guidance in hollow-core photonic crystal fiber. Opt. Express **2002**, *10*, 1195–1203.
62. Kushwaha, M.S.; Djafari-Rouhani, B. Sonic stop-bands for periodic arrays of metallic rods: Honeycomb structure. J. Sound Vib. **1998**, *218*, 697–709.
63. Economou, E.N.; Sigalas, M. Stop bands for elastic-waves in periodic composite-materials. J. Acoust. Soc. Am. **1994**, *95*, 1734–1740.
64. Marin, E.; Mangan, B.J.; Díez, A.; Russell, P.St.J. Acoustic Modes of a Dual-Core Square-Lattice Photonic Crystal Fibre Preform. In *Proceedings of the European Conference of Optics Communications*; ECOC: Amsterdam, October, 2001; 518–519.
65. Agrawal, G. *Nonlinear Fibre Optics*; Academic Press: Boston, 1989.
66. Onishi, M.; Okuno, T.; Kashiwada, T.; Ishikawa, S.; Akasaka, N.; Nishimura, M. Highly nonlinear dispersion-shifted fibers and their application to broadband wavelength converter. Opt. Fiber Technol. **1998**, *4*, 204–214.
67. Monro, T.M.; Kiang, K.M.; Lee, J.H.; Frampton, K.; Yusoff, Z.; Moore, R.; Tucknott, J.; Hewak, D.W.; Rutt, H.N.; Richardson, D.J. In *High Nonlinearity Extruded Single-Mode Holey Optical Fibers*, Optical Fiber Communication Conference (OFC'2002), Anaheim, 2002. postdeadline paper FA1-1.
68. *The Supercontinuum Laser Source*; Alfano, R.R., Ed.; Springer-Verlag: New York, 1989.
69. Ranka, J.K.; Windeler, R.S.; Stentz, A.J. Visible continuum generation in air–silica microstructure optical fibres with anomalous dispersion at 800 nm. Opt. Lett. **2000**, *25*, 25–27.
70. Wadsworth, W.J.; Ortigosa-Blanch, A.; Knight, J.C.; Birks, T.A.; Man, T.-P.M.; Russell, P.St.J.

Supercontinuum generation in photonic crystal fibres and optical fibre tapers—A novel light source. J. Opt. Soc. Am., B **2002**, *19*, 2148–2155.

71. Hartl, I.; Li, X.D.; Chudoba, C.; Ghanta, R.; Ko, T.; Fujimoto, J.G.; Ranka, J.K.; Windeler, R.S.; Stentz, A.J. Ultrahigh resolution optical coherence tomography using continuum generation in an air–silica microstructure optical fibre. Opt. Lett. **2001**, *26*, 608–610.
72. Holzwarth, R.; Reichert, J.; Udem, Th.; Hänsch, T.W.; Knight, J.C.; Wadsworth, W.J.; Russell, P.St.J. An optical frequency synthesiser for precision spectroscopy. Phys. Rev. Lett. **2000**, *85*, 2264–2267.
73. Coen, S.; Chau, A.H.L.; Leonhardt, R.; Harvey, J.D.; Knight, J.C.; Wadsworth, W.J.; Russell, P.St.J. Supercontinuum generation by stimulated Raman scattering and parametric four-wave mixing in photonic crystal fibres. J. Opt. Soc. Am., B **2002**, *19*, 753–764.
74. Harvey, J.D.; Leonhardt, R.; Wong, K.G.L.; Knight, J.C.; Wadsworth, W.J.; Russell, P.St.J. Scalar modulational instability in the normal dispersion regime using a PCF. Opt. Lett. **2003**, *28*, 2225–2227.
75. Sharping, J.E.; Fiorentino, M.; Kumar, P.; Windeler, R.S. Optical parametric oscillator based on four-wave mixing in microstructure fibre. Opt. Lett. **2002**, *27*, 1675–1677.
76. Tang, R.; Devgan, P.; Sharping, J.; Voss, P.; Lasri, J.; Kumar, P. In *Microstructure-Fibre Based Optical Parametric Amplifier in the 1550 nm Telecom Band*, Optical Fiber Communications Conference (OFC'2003), Atlanta, March 2003. paper ThT3.

Photonic Applications of Printed and Molded Nanostructures

John A. Rogers
University of Illinois at Urbana-Champagne, Urbana, Illinois, U.S.A.

INTRODUCTION

Many classes of components for integrated optics and optoelectronics rely on structures with nanometer features sizes (typically a fraction of the optical wavelength at which they operate) for couplers, filters, laser resonators, etc.[1] Similar but more advanced devices that use photonic crystals[2] as well as a variety of subwavelength elements[3,4] for bulk optics also make use of arrays of nanostructures to achieve desired optical effects. Traditionally, techniques borrowed from the microelectronics industry, such as photolithography and electron beam lithography, have been employed for these applications. The operational complexity of these methods, their high capital costs, and certain inherent limitations associated with them create research interest in alternative fabrication techniques.[5,6] Two relatively recent approaches to patterning photonic nanostructures use advanced forms of printing and molding, two of the conceptually oldest forms of lithography. These methods offer submicron resolution along with patterning capabilities that are well suited to this class of application. This article summarizes these methods and demonstrates their use in fabricating a range of components for integrated optics.

HIGH-RESOLUTION STAMPS AND MOLDS

High-resolution stamps and molds can be formed with electron beam lithography or photolithography according to the procedures illustrated in Fig. 1. Casting and curing a polymer against a pattern of resist formed with these techniques (known as the "master") yields polymeric elements that can serve as stamps and molds. Many elements can be produced from a single master and each element can be used many times. Similar molds and stamps can be made out of the material that serves as the substrate for the master by using the resist to generate surface relief, through selective etching. There are currently two classes of printing techniques that use these stamps for patterning with submicron resolution—microcontact printing (μCP)[7,8] and nanotransfer printing (nTP).[9,10] In μCP, an elastomeric stamp [usually poly(dimethylsiloxane), PDMS; Dow Corning, Inc.] is used to print patterns of chemical inks. In nTP, a hard or an elastomeric stamp delivers a solid ink material to a substrate. μCP offers typical edge resolution of 50–200 nm depending on the inks and substrates. Demonstrated resolution in nTP is in the range of 5–20 nm. The stamps for these techniques can also be used as high resolution molds for embossing relief structures into polymers and other soft materials.[5]

Microcontact Printing

It was mainly developed for use with inks that form self-assembled monolayers (SAMs) of alkanethiolates on gold and silver. The procedures for carrying out μCP in these systems are simple: A stamp, inked with a solution of alkanethiol, is brought into contact with the surface of a substrate to transfer ink molecules to regions where the stamp and substrate contact. The resolution and effectiveness of this method relies on conformal contact between the stamp and the surface of the substrate, rapid formation of highly ordered monolayers,[11] and autophobicity of the SAM,[12] which effectively blocks the reactive spreading of the ink across the surface. It can pattern SAMs over relatively large areas (up to $\sim$0.25 ft^2 have been demonstrated in prototype electronic devices)[13] in a single impression.

Fig. 2 schematically illustrates the use of μCP and wet etching to pattern a thin film of Au. Fig. 3 shows scanning electron microscope (SEM) images of representative nanopatterns of gold (20-nm thick, thermally evaporated with a 2.5-nm layer of Ti as an adhesion promoter) that were fabricated on silicon wafers using this approach. In both of these examples, the masters for the stamps consist of photoresist patterned on silicon wafers with projection-mode photolithography. Placing a few drops of a $\sim$1 mM solution of hexadecanethiol (HDT) in ethanol on the surface of the stamps, and then blowing them dry with a stream of nitrogen prepares them for printing. Contacting the metal film for a few seconds with the stamp produces a patterned self-assembled monolayer (SAM) of HDT. An aqueous etchant removes the unprinted regions of the gold.[14] The patterns in Fig. 3 have edge roughness of

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009349

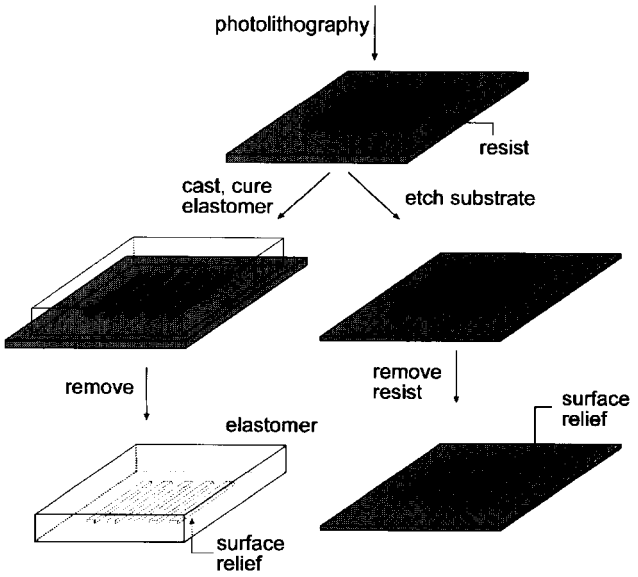


Fig. 1 Schematic illustration of two representative methods for producing high-resolution stamps and molds. Both begin with photolithography to pattern a thin layer of resist on a flat substrate, such as a silicon wafer. This structure, known as the master, is converted to a stamp or mold either by etching or by molding. In the first case, the resist acts as a mask for etching the underlying substrate. Removing the resist yields a stamp or a mold. In the molding approach, a prepolymer is cast against the relief structure formed by the patterned resist on the substrate. Curing (thermally or optically) and then peeling the resulting solid polymer away from the substrate completes the process. (*View this art in color at www.dekker.com.*)

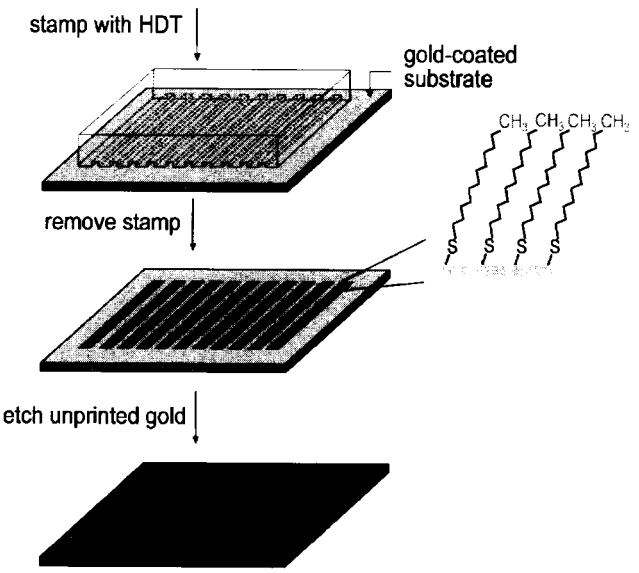


Fig. 2 Schematic illustration of an example of microcontact printing. The first step involves "inking" a "stamp" with a solution of hexadecanethiol (HDT) in ethanol. Contacting the stamp to a thin film of gold on a substrate produces a patterned SAM of HDT. The printed SAM can act as a resist for the aqueous-based wet etching of the exposed regions of the Au. The resulting pattern of conducting gold can be used to build devices of various types. (*View this art in color at www.dekker.com.*)

~50–100 nm. This resolution, which is typical for μCP, is determined by the grain size of the metal films, the isotropic etching process, slight reactive spreading of the inks and edge disorder in the patterned SAMs. Patterns with these feature sizes are useful for many applications in integrated optics, and especially those at the infrared wavelengths (1.5 μm) that are predominant in communication systems.

Nanotransfer Printing

Nanotransfer printing (nTP) is a more recent high-resolution printing technique that uses surface chemistries as interfacial "glues" and "release" layers to control the transfer of solid material layers from relief features on a stamp to a substrate.[9,10] This approach, which is illustrated in Fig. 4, is purely additive (i.e., material is only deposited in locations where it is needed) and it can generate complex patterns of single or multiple layers of materials with nanometer resolution over large areas in a single process step. It does not suffer from surface diffusion or edge disorder in the patterned "inks" of μCP nor does it require postprinting etching or deposition steps to produce structures of functional materials. The method involves four components: 1) a stamp (rigid, flexible, or elastomeric) with relief features in the geometry of the desired pattern; 2) a method for depositing a thin layer of solid material onto the raised features of this stamp; 3) a means for bringing the stamp into intimate physical contact with a substrate; and 4) surface chemistries that prevent adhesion of the deposited material to the stamp and promote its strong adhesion to the substrate. nTP has

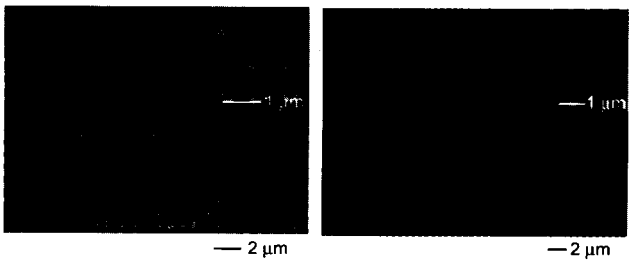


Fig. 3 Scanning electron micrographs of typical structures formed by microcontact printing a self-assembled monolayer ink of hexadecanethiol onto thin gold films followed by etching of the unprinted areas. The left frame shows an array of Au (20-nm thick) dots with ~500-nm diameters. The right frame shows printed lines of Au (20-nm thick). The edge resolution of features in both of these patterns is ~50–100 nm.

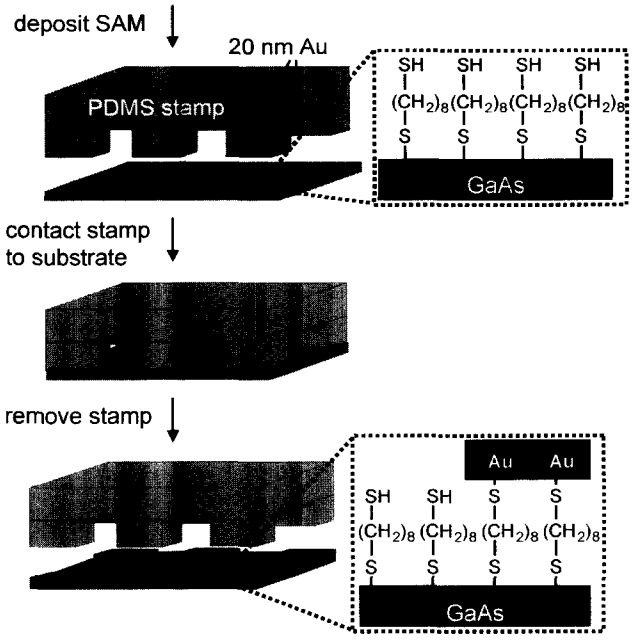


Fig. 4 Schematic illustration of procedures for nanotransfer printing a pattern of Au on a GaAs substrate. A uniform monolayer of 1,8-octanedithiol molecules is formed on the GaAs using vapor or solution phase deposition. Evaporating a collimated flux of Au (20 nm) onto an elastomeric stamp forms layers on the recessed and raised regions but not on the sidewalls. Contacting this stamp with the substrate leads to Au–S bonds. Removing the stamp leaves a pattern of gold on the substrate in the geometry of relief on the stamp. The transferred gold patterns are effectively "glued" to the surface by strong Ga–S, As–S, and S–Au bonds at the interface. The insets highlight the idealized orientation of the dithiol molecule on GaAs surface; control experiments indicate that the unreacted thiol endgroups are generally oriented away from the wafer surface.

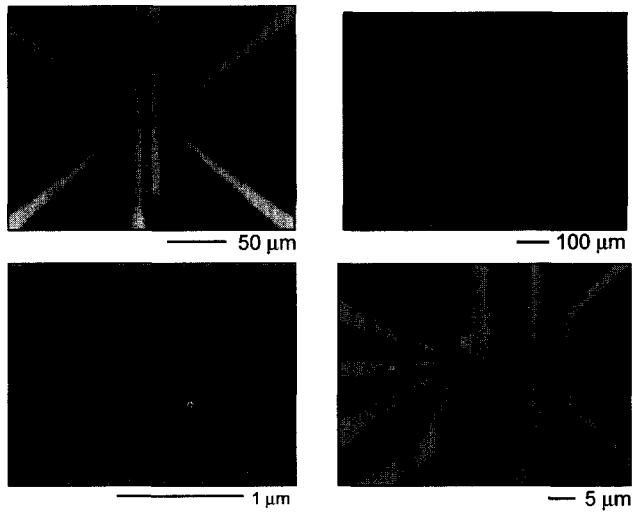


Fig. 5 Optical micrographs of patterns of Au (15-nm thick) formed on plastic (top left and bottom left frames), silicon (top right frame), and GaAs (bottom right frame) substrates with nanotransfer printing. The transfer chemistry in the top left case relies on silane self-assembled monolayers with exposed thiol groups. The chemistry for the patterns in the top and bottom right cases relies on surface dehydration reactions. The bottom right pattern used a dithiol chemistry. The minimum feature sizes and the edge resolution are limited, in all cases, by the stamps. (*View this art in color at www.dekker.com.*)

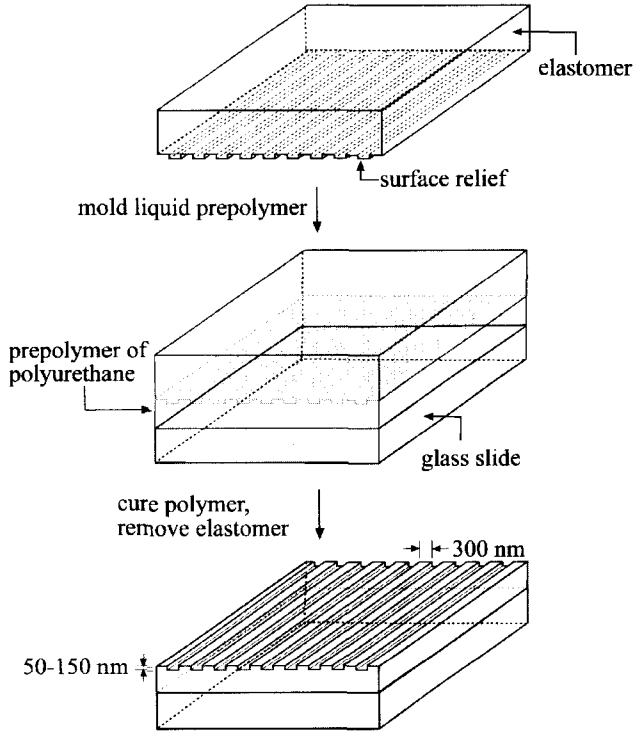


Fig. 6 Schematic illustration of molding procedures for defining distributed feedback gratings for lasers. The gratings are formed by exposing a thin film of a photocurable polymer to ultraviolet light through an elastomeric mold.

been demonstrated with SAMs and other surface chemistries for printing onto flexible and rigid substrates with hard inorganic and soft polymer stamps.[9,10,15] Fig. 4 presents a set of procedures for using nTP to pattern a thin layer of Au with a surface transfer chemistry that relies on Au–S bonding and self-assembled monolayers.[15] The process begins with fabrication of rigid or elastomeric stamps using the procedures of Fig. 1. Careful control of the processing steps yields features of relief with nearly vertical or slightly reentrant sidewalls. Directional deposition of Au (or other materials) onto the surfaces of such stamps yields coatings only on the raised and recessed regions and not on the sidewalls. The depth of relief on the stamps is typically >200 nm for patterning metal films with thicknesses <50 nm.

In the case of Fig. 4, contacting a gold-coated PDMS stamp to a GaAs substrate that has a self-assembled dithiol monolayer on its surface leads to thiol-gold linkages at the interface. Removing the stamp (to which the gold does not

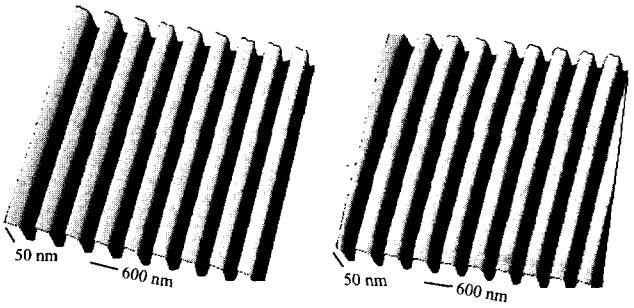


Fig. 7 Atomic force micrographs of distributed feedback grating structures. The right frame was collected from a photolithographically defined pattern in a thin layer of SiO_2 on a silicon wafer. The left frame was collected from a polymer replica generated with an elastomeric mold produced from the structure illustrated on the right.

adhere) leaves behind a pattern of gold in the geometry of the raised regions of the stamp.[15] Similar procedures can be used with a wide range of stamps, patterned materials, substrates, and surface chemistries. Fig. 5 shows some representative examples. The resolution of nTP is limited mainly by the roughness at the edges of the films that are deposited on the stamps and by the stamps themselves. This edge roughness is less than 5–20 nm for evaporated gold and stamps fabricated by electron beam lithography. There is no observable degradation in feature sharpness or resolution because of the transfer process. Care must be taken to use optimized deposition and printing conditions in order to avoid nanocracking in the transferred layer, especially when patterns of rigid materials are printed from deformable stamps onto deformable surfaces. In the applications that we have explored, even when they might exist, the nanocracks (which are not visible with conventional optical or scanning electron microscopes) do not significantly affect the functional properties of the devices

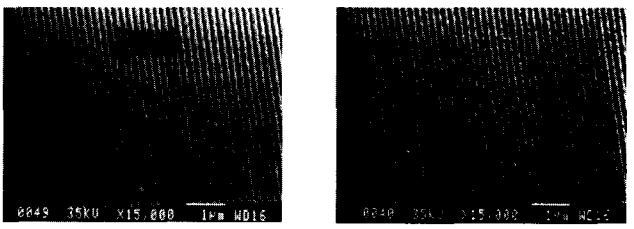


Fig. 8 Scanning electron micrographs of 200 nm period distributed feedback gratings in a thin layer of SiO_2 on a silicon wafer (left frame) and in a thin layer of a photocurable polymer on a glass slide (right frame). The structure in silica was formed by direct write electron beam lithography and etching. The structure in the photopolymer was defined by molding using an elastomeric mold formed by casting and curing against the silica structure.

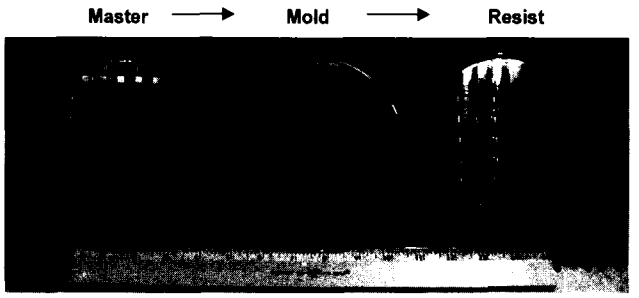


Fig. 9 Images of a large-area master (6-in. silicon wafer with patterned photoresist; left frame), a large-area mold formed from this master (middle frame), and a molded thin film (2 μm thick) of photopolymer formed using this mold (right frame). The mechanical flexibility and elastomeric nature of the mold are key features that allow this type of large-area replication. The lateral feature sizes in this pattern range from 200 nm to 1 cm. The relief depth ranges from 1 to 0.25 μm. (*View this art in color at www.dekker.com.*)

(e.g., resistivity test structures, organic transistors and simple circuits, photonic structures, thin film capacitors, etc.).

Nanomolding

The stamps used for μCP and nTP can also serve as molds for generating surface relief features in other materials. Nanoimprint lithography and related embossing methods represent one class of such techniques that use hard stamps and thermally softened polymers or photocurable liquids.[5] The use of elastomeric materials for "soft" molds yields a related type of technique which, because of the flexibility of the molding elements, can be used for

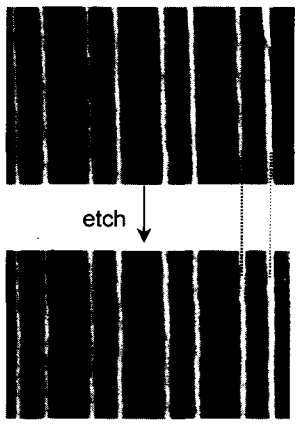


Fig. 10 Scanning electron micrographs showing molded features of relief in a thin polymer film before (top frame) and after (bottom frame) reactive ion etching to remove the thin regions.

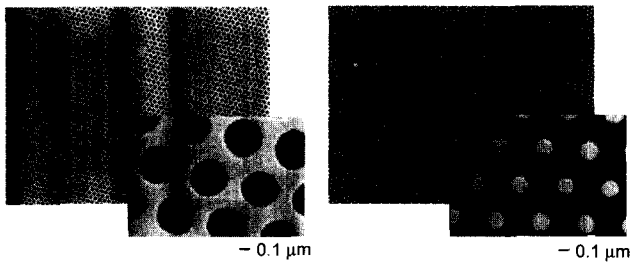


Fig. 11 Scanning electron micrographs of nanopatterns of tungsten (left frame) and gold (right frame) formed by etching through a sacrificial polymer resist defined by molding and etching.

large-area patterning of nanostructures even in fragile materials.[6] Fig. 6 shows procedures for using this approach to generate features of relief in a thin photocurable polymer. The transparent PDMS mold allows the photopolymer to be cured by exposure to ultraviolet light passing through the mold. In the case illustrated here, the pattern consists of a distributed feedback grating for a laser that emits in the visible spectrum. Fig. 7 presents atomic force microscope images of a relief pattern generated by photolithography and etching of a thin layer of SiO_2 on a silicon substrate (left frame) and a fifth generation replica grating (right frame) in a photocurable polyurethane (NOA 72, Norland) formed by using the procedures of Fig. 6 with a PDMS mold generated from the SiO_2 structure.[16] These procedures are effective for pattern replication down to the 50–100-nm scale. Fig. 8 shows scanning electron micrographs of 100 nm lines and spaces fabricated by electron beam lithography and etching in SiO_2 (left frame) and a replica grating structure in a molded polymer (right frame).[17] Below these sizes, the mechanical properties of the mold can be limiting;[18–21] newer elastomers with comparatively high modulus and physical toughness can extend the resolution of these procedures deep into the sub-100-nm regime.[18,20]

The resolution and replication fidelity illustrated in Figs. 7 and 8 can be achieved over large areas. Fig. 9 shows a master, a mold, and a replicated pattern in a thin polymer film carried out with and on 6-in. silicon wafers. The ability to flex and bend the mold allows contact with the thin spin cast photopolymer to be initiated at the center of the wafer. As the contact line is gradually allowed to move toward the edges of the substrate, air is pushed out, thereby eliminating trapped air pockets of the type that are difficult to avoid in more conventional embossing and imprinting techniques that use rigid molds. The mechanical flexibility of the mold is also important for release because it allows the mold to be removed from the substrate by bending and peeling, in way that avoids damage to the substrate or the mold.

The wafer-scale results illustrated in Fig. 9 incorporate features with lateral dimensions down to 200 nm. This relief can be converted into isolated polymer nanostructures by etching to remove the thin regions of the molded films. Fig. 10 shows SEMs of a molded thin polymer film before and after exposure to an oxygen reactive ion etch. Careful control of the etching conditions can enable removal of the thin regions of the molded films without significant changes in the lateral dimensions of the features. The isolated polymer lines that are generated by these procedures can be used in the same way that patterned photoresist is used in conventional semiconductor processing. In the examples illustrated in Fig. 11, similar polymer structures were used for dry etching a pattern into underlying films of tungsten and gold. This molding and etching approach provides an alternative route to building the types of structures that can be formed by μCP and nTP.

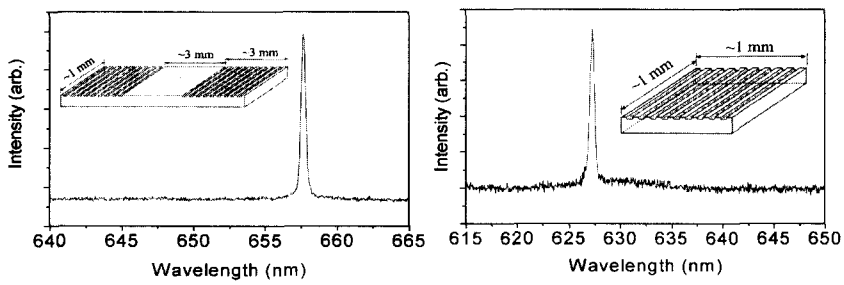


Fig. 12 Schematic illustrations and emission spectra of plastic lasers that use molded resonators based on surface relief distributed Bragg reflectors (DBRs) and distributed feedback gratings (DFBs) on glass substrates. The grating periods are ~600 nm in both cases. The lasers use thin-film plastic gain media deposited onto the printed gratings. The emission profiles, the lasing thresholds, and other characteristics of these devices are comparable to similar lasers that use resonators formed by high-resolution projection-mode photolithography.

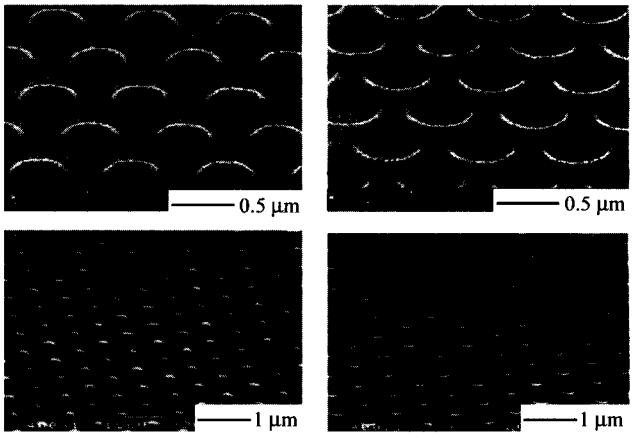


Fig. 13 Scanning electron micrographs of four photonic crystal structures (~50-nm deep) molded onto the surfaces of sol–gel-derived glass films. [The slight "graininess" in these images is due to a thin coating of gold (~5 nm) deposited to eliminate charging.]

APPLICATIONS

Nanopatterns such as those described in the preceding discussion can be used for a variety of photonic devices. To build lasers out of molded or printed surface relief structures (which not only yields functional devices but also demonstrates the quality of these structures—defect density, spatial coherence, uniformity, etc.—for other applications), a thin layer of organic gain media is deposited on top of the structure to define a planar waveguide that confines light to the surface. The thickness of this layer is chosen to support a single transverse mode. Photopumping the structure generates lasing as a result of feedback associated with reflections from the grating. A lower-order interaction couples light out of the waveguide. For the examples described in this article, a third-order interaction with the grating provides the feedback and a first-order interaction leads to output coupling.

Fig. 12 shows the typical performance of plastic lasers that use printed and molded distributed feedback (DFB) and distributed Bragg (DBR) resonators with spin-coated thin film (150–200-nm thick) gain media of 2-(4-biphenyl)-5-(4-*tert*-butylphenyl)-1,3,4-oxadiazole doped with 1 wt.% Coumarin 490 and DCMII (Exciton, Inc.).[16] Photopumping these structures with 2-nsec pulses from a nitrogen laser with intensities >5 kW/cm^2 leads to multimode lasing at resolution-limited linewidths and wavelengths corresponding to the third harmonic of the gratings. The spectra and the threshold intensities for lasing are similar to those observed in lasers that use SiO_2 gratings generated with photolithography. They are better than those that we observed from gratings formed by pressing heated thin polymer films against rigid templates.[22]

The same printing and molding techniques can be used with more unusual materials and laser structures. Fig. 13 shows images of some two-dimensional photonic crystal structures molded into a type of sol–gel derived glass[23] as examples. This class of material, known as an organically modified silicate, consists of interpenetrating organic and inorganic networks prepared by using a metal alkoxide with a reactive organic moiety. This hybrid matrix forms, at relatively low temperatures, thick, mechanically hard coatings with good optical properties. To generate fine patterns of surface relief in these films, the curing is carried out while the film is in contact with a PDMS mold.[23] The structures in Fig. 13 are molded in

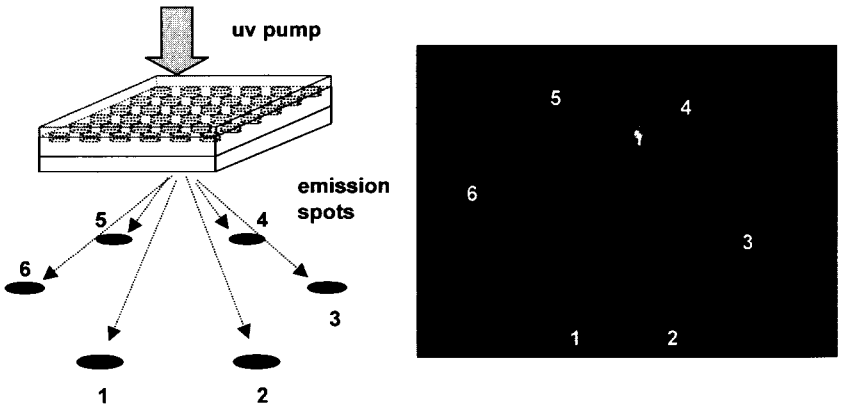


Fig. 14 Schematic illustration (left frame) and optical image (right frame) of emission from a photopumped photonic crystal laser that uses a molded resonator of cylindrical depressions ~50-nm deep, with 400-nm diameters and center-to-center separations of 600 nm in a layer of an organically modified silicate glass film on glass slide. In this device, the crystal acts both as the resonator to provide feedback for the laser and the output coupler. (*View this art in color at www.dekker.com.*)

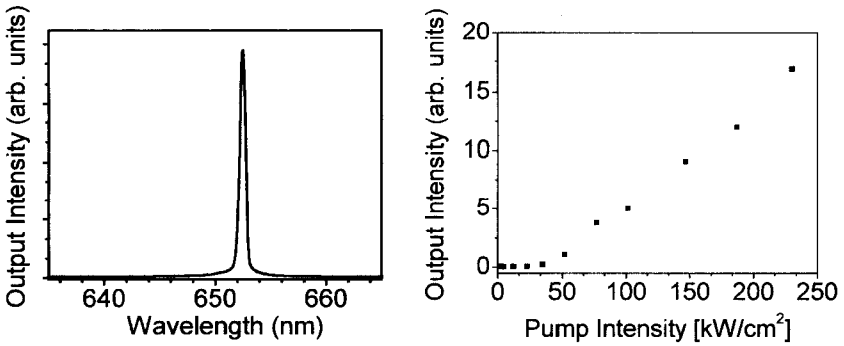


Fig. 15 Emission spectrum (left frame) and output intensity as a function of pump intensity (right frame) for a photopumped laser that uses a molded two-dimensional photonic crystal as a resonator. The narrow emission linewidth (instrument limited), the well-defined emission spots and the relatively low lasing threshold provide evidence that the quality of the molded resonator is high (i.e., spatially uniform and free of defects).

organically modified silicate films several microns thick and consist of periodic arrays of cylindrical posts or holes with submicron diameters, arrayed in triangular, square, and honeycomb lattices. The images suggest that the structures are replicated with a spatial resolution that comfortably exceeds both the smallest dimensions of the features (~150 nm) in these sol–gel-derived photonic crystals and in structures (~250 nm) that are useful for applications in telecommunications.

Depositing thin films (150–200-nm thick) of organic gain materials [8-hydroxyquinolinato aluminum (Alq) and the laser dye, DCMII] onto these molded patterns produces lasers that operate due to Bragg reflections induced by the photonic lattice.[23] Fig. 14 illustrates the photopumping geometry (left frame) and an image of the spatial pattern of laser emission (right frame). A typical spectrum (left frame) and the measured dependence of the output intensity on the intensity of the pump light (right

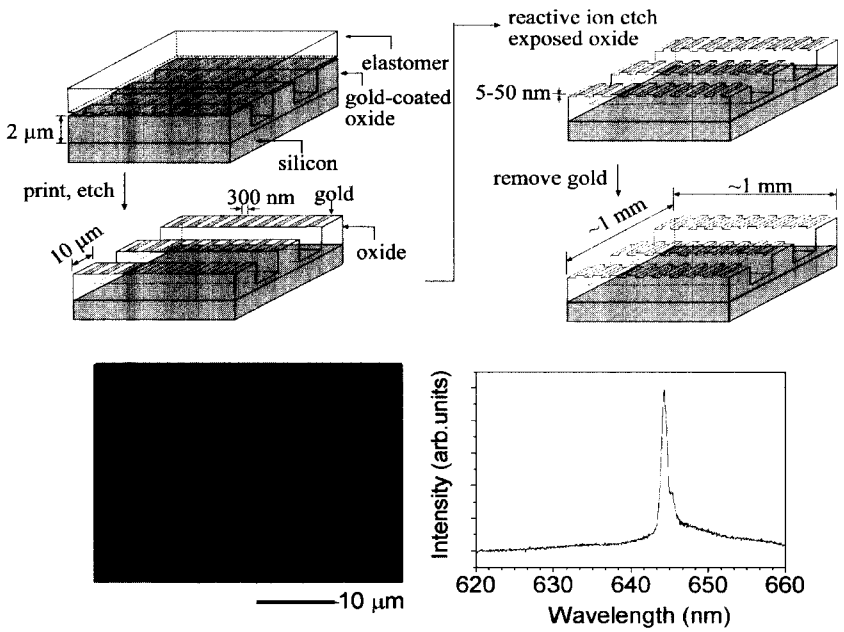


Fig. 16 Schematic illustration (top frames) of steps for microcontact printing high-resolution gratings directly onto the top surfaces of ridge waveguides. The printing defines a sacrificial etch mask of gold, which is subsequently removed. The bottom left frame shows a top-view optical micrograph of printed gold lines on the ridge waveguides. The bottom right frame shows the emission output of a plastic photopumped laser that uses the printed structure and a thin evaporated layer of gain media. (*View this art in color at www.dekker.com.*)

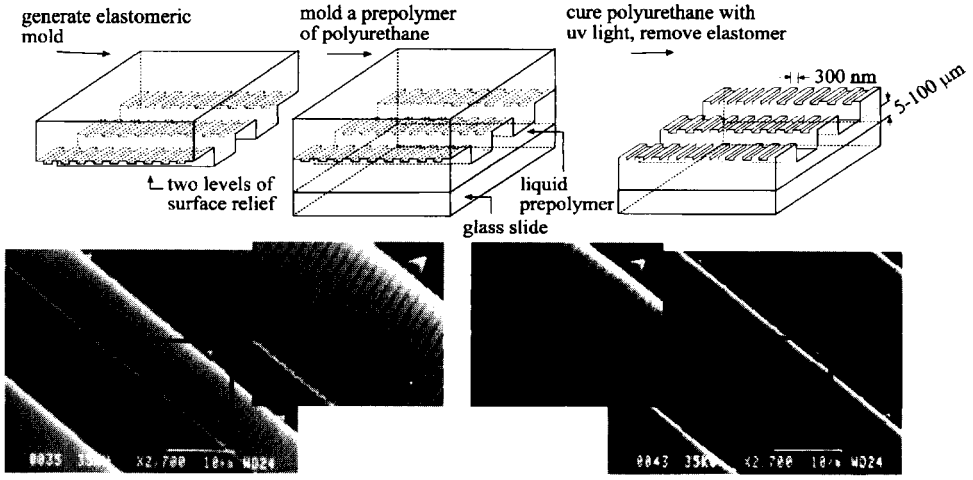


Fig. 17 Schematic illustration (top frames) of the use of a two-level mold to produce, in a single step, distributed feedback ridge waveguide laser resonators. The scanning electron micrographs on the right show the silica master that was used for the mold (top frame) and a molded replica in a photocurable polymer (bottom frame).

frame) are shown in Fig. 15. The lasing threshold ($\sim$50 kW/cm^2) and the emission spectra are, to within device-to-device variations, identical to those of similar lasers that use photolithographically defined resonators in SiO_2.[24]

Printing and molding procedures not only provide a simple, low-cost route to structures that can be fabricated with other approaches, but they also enable some unique patterning capabilities. For example, μCP can be used to form DFB resonators directly on the top surfaces ridge waveguides.[25] In certain cases, there are manufacturing advantages to forming the gratings after defining the waveguides. Achieving this process sequence with photolithography is challenging because spin cast layers of photoresist develop large and undesired thickness variations due to dewetting at the sharp edges of relief associated with the waveguides. Fig. 16 illustrates the procedures for using μCP for this patterning task.[25] The

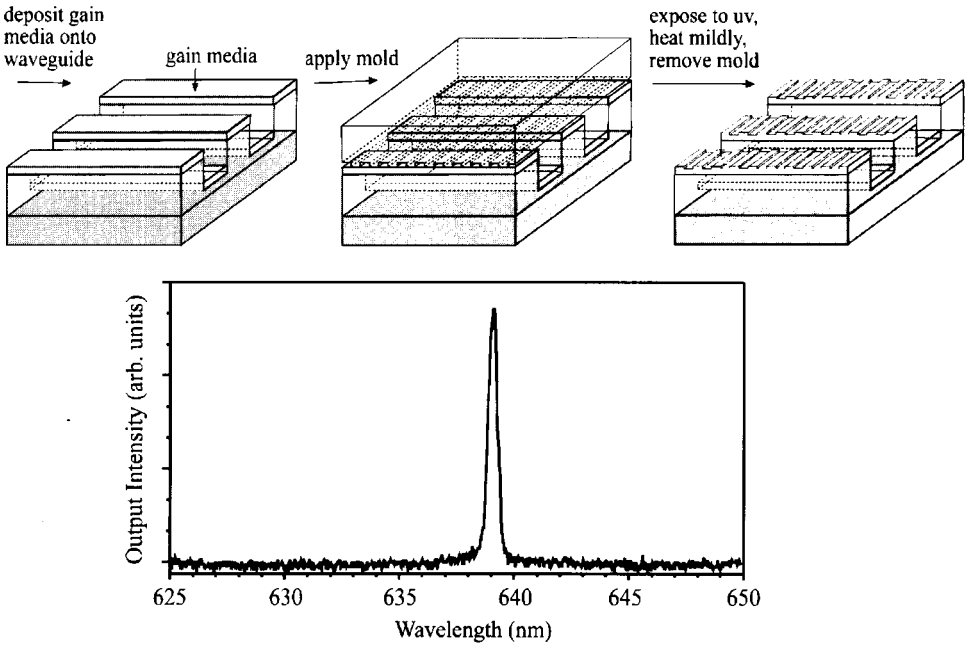


Fig. 18 Schematic illustration (top frames) of procedures for molding high-resolution gratings directly onto the surface of a thin layer or organic gain material. Photopumping this structure leads to lasing. The bottom frame shows a typical emission spectra.

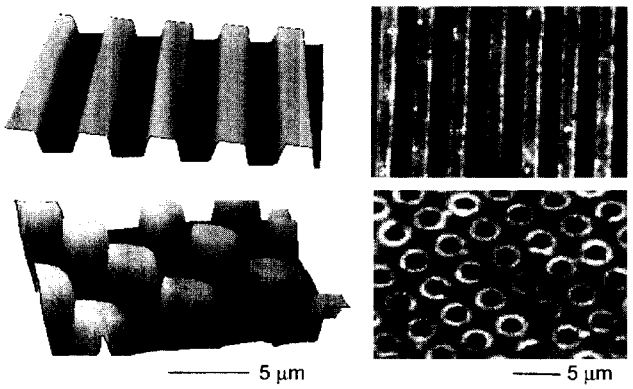


Fig. 19 Atomic force micrographs (left frames) of surface relief structure molded into a thin film of poly(*p*-phenylene vinylene) on a thin film of indium tin oxide on a glass slide. The right frames show optical micrographs of the patterns of emission that results from diodes that are formed by depositing Al/Ca on top of the molded films.

bottom left frame shows an optical micrograph of the printed gold lines. Sublimation of a ~200-nm film of Alq doped with 0.5–5.0 wt.% DCMII onto the resonators produces waveguide DFB lasers. The bottom right frame of Fig. 16 shows the emission profile. Once the ridge waveguide DFB structure has been fabricated, it can be used as a master for producing a mold with two levels of relief.[17] The structure can then be rapidly replicated by molding in a single step. Fig. 17 shows the procedures and micrographs of the master and a typical replica in a photopolymer.

This basic type of laser can also be produced by molding the relief structure directly into the organic gain material.[25] Fig. 18 shows the procedures. In this case, a PDMS mold is placed in contact with a thin layer of gain material that is evaporated onto the waveguides. Exposing this film to ultraviolet laser light while in contact with the mold softens the material and causes it to conform to the relief of the mold. Removing the PDMS mold after this procedure leaves a surface relief DFB grating embossed onto the gain material itself. Photopumping this structure leads to lasing; a typical spectrum is shown in Fig. 18. We have also used this same basic soft molding approach to pattern relief features into thin films of the polymer poly(*p*-phenylene vinylene) (PPV) to structure light emission in polymer light-emitting diodes (Fig. 19). A PPV film cast on top of a piece of glass coated with indium tin oxide was molded during solvent evaporation using a PDMS mold. A uniform top electrode consisted of a uniform bilayer of Al/Ca. Here thickness relief produced by the molding leads to light output from only the thin regions of the film at certain voltages.[26] This sort of direct patterning of thin layers of electroactive organics could have important applications in the emerging field of active organic optoelectronics. These classes of mechanically and chemically fragile materials are difficult or impossible to pattern with conventional approaches because of their chemical incompatibility with the required photoresists, developers, and solvents.

CONCLUSION

This article summarizes some of our recent work in the development and application of unconventional methods for fabricating photonic nanostructures, with examples of their use in organic optoelectronics. Integrated optics represents a field where new or emerging techniques for nanofabrication could find important niche applications. The dimensional requirements for registration and alignment (2–500 μm typically) in these systems is much less demanding than the requirements on the minimum feature sizes (~100–300 nm for the devices described here). As a result, the complex and highly engineered systems that are required for nanoscale registration are not needed, and the challenges associated with achieving nanoscale overlay accuracy with deformable molds and stamps can be avoided.[27] These features greatly simplify the potential introduction of these new lithographic methods into realistic manufacturing settings. Compared to other types of unconventional approaches that have nanometer resolution (e.g., imprint lithography, embossing, etc.) the printing and molding methods described here are attractive, in part because they are directly compatible with classes of organic active materials. The technologies that emerge from the combined use of soft materials and soft patterning methods for photonics have a strong potential to yield important classes of devices for future optical communication networks, information displays, and other systems.

ACKNOWLEDGMENTS

The author would like to thank all of the collaborators that contributed to the work described here.

REFERENCES

1. *Integrated Optics,* 2nd Ed.; Tamir, T., Ed.; Springer-Verlag, 1985.
2. Yablonovich, E. Photonic crystals: Semiconductors of light. Sci. Am. December **2001**, 47–55 and references therein.
3. Clapham, P.B.; Hutley, M.C. Reduction of lens reflexion by the 'moth eye' principle. Nature (Lond.) **1973**, *244*, 281–282.

4. Enger, R.C.; Case, S.K. Optical elements with ultrahigh spatial-frequency surface corrugations. Appl. Opt. **1983**, *22*, 3220–3228.
5. Mirkin, C.A.; Rogers, J.A. Emerging methods for micro- and nanofabrication. MRS Bull. **2001**, *26*, 506–507. and references therein.
6. Xia, Y.; Rogers, J.A.; Paul, K.E.; Whitesides, G.M. Unconventional methods for fabricating and patterning nanostructures. Chem. Rev. **1999**, *99* (7), 1823–1848.
7. Kumar, A.; Whitesides, G.M. Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol ink followed by chemical etching. Appl. Phys. Lett. **1993**, *63*, 2002–2004.
8. Michel, B.; Bernard, A.; Bietsch, A.; Delamarche, E.; Geissler, M.; Juncker, D.; Kind, H.; Renault, J.P.; Rothuizen, H.; Schmid, H.; Schmidt-Winkel, P.; Stutz, R.; Wolf, H. Printing meets lithography: Soft approaches to high-resolution printing. IBM J. Res. Develop. **2001**, *45*, 697–719.
9. Loo, Y.-L.; Willett, R.W.; Baldwin, K.; Rogers, J.A. Interfacial chemistries for nanoscale transfer printing. J. Am. Chem. Soc. **2002**, *124* (26), 7654–7655.
10. Loo, Y.-L.; Willett, R.W.; Baldwin, K.; Rogers, J.A. Additive, nanoscale patterning of metal films with a stamp and a surface chemistry mediated transfer process: Applications in plastic electronics. Appl. Phys. Lett. **2002**, *81* (3), 562–564.
11. Larsen, N.B.; Biebuyck, H.; Delamarche, E.; Michel, B. Order in microcontact printed self-assembled monolayers. J. Am. Chem. Soc. **1997**, *119*, 3017–3026.
12. Biebuyck, H.A.; Whitesides, G.M. Self-organization of organic liquids on patterned self-assembled monolayers of alkanethiolates on gold. Langmuir **1994**, *10*, 2790–2793.
13. Rogers, J.A.; Bao, Z.; Baldwin, K.; Dodabalapur, A.; Crone, B.; Raju, V.R.; Kuck, V.; Katz, H.; Amundson, K.; Ewing, J.; Drzaic, P. Paper-like electronic displays: Large area, rubber stamped plastic sheets of electronics and electrophoretic inks. Proc. Natl. Acad. Sci. **2001**, *98* (9), 4835–4840.
14. Xia, Y.N.; Zhao, X.M.; Kim, E.; Whitesides, G.M. A selective etching solution for use with patterned self-assembled monolayers of alkanethiolates on gold. Chem. Mater. **1995**, *7*, 2332–2337.
15. Loo, Y.-L.; Hsu, J.W.P.; Willett, R.L.; Baldwin, K.W.; West, K.W.; Rogers, J.A. High-resolution transfer printing on GaAs surfaces using alkane dithiol self-assembled monolayers. J. Vac. Sci. Technol. B **2002**, *20*, 2853–2856.
16. Rogers, J.A.; Meier, M.; Dodabalapur, A. Using stamping and molding techniques to produce distributed feedback and Bragg reflector resonators for plastic lasers. Appl. Phys. Lett. **1998**, *73* (13), 1766–1768.
17. Rogers, J.A.; Bao, Z.; Dodabalapur, A.; Schueller, O.J.A.; Whitesides, G.M. Printing, molding and near-field photolithographic methods for patterning organic lasers, smart pixels and simple circuits. Synth. Met. **2000**, *115* (1–3), 5–11.
18. Schmid, H.; Michel, B. Siloxane polymers for high-resolution, high-accuracy soft lithography. Macromolecules **2000**, *33*, 3042–3049.
19. Odom, T.W.; Love, J.C.; Wolfe, D.B.; Paul, K.E.; Whitesides, G.M. Improved pattern transfer in soft lithography using composite stamps. Langmuir **2002**, *18*, 5314–5320.
20. Choi, K.M.; Rogers, J.A. A photocurable poly-(dimethylsiloxane) for soft lithography in the nanometer regime. J. Am. Chem. Soc. **2003**, *125*, 4060–4061.
21. Hui, C.Y.; Jagota, A.; Lin, Y.Y.; Kramer, E.J. Constraints on microcontact printing imposed by mechanical deformation. Langmuir **2002**, *18* (4), 1394–1407.
22. Berggren, M.; Dodabalapur, A.; Slusher, R.E.; Timko, A.; Nalamasu, O. Organic solid-state lasers with imprinted gratings on plastic substrates. Appl. Phys. Lett. **1998**, *72*, 410–411.
23. Schueller, O.J.A.; Whitesides, G.M.; Rogers, J.A.; Meier, M.; Dodabalapur, A. Fabrication of photonic crystal lasers by nanomolding of sol–gel glasses. Appl. Opt. **1999**, *38* (27), 5799–5802.
24. Berggren, M.; Dodabalapur, A.; Slusher, R.E.; Bao, Z.; Timko, A.; Nalamasu, O. Electron. Lett. **1998**, *34*, 90–91.
25. Rogers, J.A.; Meier, M.; Dodabalapur, A. Distributed feedback ridge waveguide lasers fabricated by nanoscale printing and molding on non-planar substrates. Appl. Phys. Lett. **1999**, *74* (22), 3257–3259.
26. Rogers, J.A.; Bao, Z.; Dhar, L. Fabrication of patterned electroluminescent polymers that emit in geometries with features into the sub-micron range. Appl. Phys. Lett. **1998**, *73* (3), 294–296.
27. Rogers, J.A.; Paul, K.E.; Whitesides, G.M. Quantifying distortions in soft lithography. J. Vac. Sci. Technol., B **1998**, *16*, 88–97.

Photovoltaics for the Next Generation: Organic-Based Solar Cells

Sean E. Shaheen
David S. Ginley
National Renewable Energy Laboratory (NREL), Golden, Colorado, U.S.A.

INTRODUCTION

The advancement of renewable energy technology has become increasingly critical as an environmentally sustainable path to world economic and societal development. Large-scale harvesting of energy from sunlight using photovoltaic devices offers tremendous potential as a renewable energy resource. Although the field of photovoltaics has continually evolved over the last several decades, many researchers are looking beyond current device designs to what have been termed third-generation technologies. These technologies are aimed at producing significant improvements in device efficiency-to-cost ratios. This necessitates significantly improving efficiency or reducing cost or ideally both. To realize these goals, many of these technologies will need to utilize nanostructured materials and composite systems that can be tailored to have optimized electronic and optical properties.

Organic/polymeric molecule-based electronics are a promising route to meeting these needs. Organic semiconductors have emerged as a class of materials that can be specifically designed to have a wide range of chemical, optical, and electronic properties, yet they can be processed via low-cost, solution-based techniques. The field has now matured to a point where most of the fundamental photophysical and electronic phenomena of these materials are reasonably well understood. The current challenge is to fully utilize the unique properties of organic semiconductors and to specifically design and fabricate materials and devices that lead to high efficiencies. One critical aspect of these materials is the molecular morphology. It has become apparent that the particular structural arrangement of molecules in an organic semiconductor device can have a large impact on properties such as charge carrier mobility and exciton diffusion. Short-range ordering of the molecules is primarily controlled by intermolecular interactions, as well as by processing conditions such as choice of solvent. Learning how to control the morphology on longer length scales will be a necessity to produce optimized device structures.

Parallel to the development of organic semiconductors has been progress in low-cost synthetic techniques for metal oxide semiconductors. A variety of solution-based and electrochemical routes have been developed that result in materials with well-defined, novel nanostructures that can be easily controlled by varying the fabrication conditions. This then opens up the possibility of low-cost fabrication of organic–inorganic composite structures, in which the long-range morphology of the organic molecules is determined by the three-dimensional structure of the metal oxide semiconductor. Ultimately, the development of efficient photovoltaic devices will require precise control over many parameters, including what wavelengths of light are absorbed to create excitons, the pathways available for exciton migration, the pathways and mobilities for charge carrier transport, and the alignment of energetic levels at interfaces. This article will review existing organic photovoltaic devices and explore initial efforts at more sophisticated, yet low-cost, organic–inorganic composite structures aimed at achieving high power conversion efficiencies.

TECHNOLOGIES FOR PHOTOVOLTAIC POWER CONVERSION

In 2001, 38.6 quads (1 quad=1 quadrillion Btu= 1.055×10^9 GJ) of energy, largely from fossil fuels, was consumed to produce electricity in the United States. After conversion losses, 12.7 quads of net electrical energy was output by power plants for general consumption.[1] To produce this amount of energy in 1 year would require covering a 100×100 km square area of the Nevada desert with solar cells with a power conversion efficiency of 15%. Thus in order for photovoltaics to make a significant impact on energy production strategy, they must be produced in very large areas at low cost.

Worldwide production levels for solar cell modules have been growing rapidly over the last several years, with Japan recently taking the lead in total production volume (Table 1). Current production is dominated by single crystal and polycrystalline silicon modules, which represent a little over 82% of the market. These devices, based on silicon wafers, have been termed the "first generation"

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009307

Table 1 World photovoltaic module production, consumer and commercial (MW)

Country	1994	1995	1996	1997	1998	1999	2000	2001
Japan	16.5	16.4	21.2	35.0	49.0	80.0	128.6	171.22
United States	25.64	34.75	38.85	51.0	53.7	60.8	74.97	100.32
Europe	21.7	20.1	18.8	30.4	33.5	40.0	60.66	86.38
ROW	5.6	6.35	9.75	9.4	18.7	20.5	23.42	32.62
Total	*69.44*	*77.6*	*88.6*	*125.8*	*154.9*	*203.1*	*287.65*	*390.54*

ROW=rest of the world.
Source: PV News, Vol. 21, No.2, Feb. 2002.

of photovoltaic technology. These are single-junction devices, which are limited by thermodynamic considerations to a maximum theoretical power conversion efficiency of 31% under direct AM1.5 sunlight.[2] Fig. 1 shows the progress over the last two decades in single crystal silicon laboratory devices, which have achieved efficiencies close to 25%.[3] Commercially produced large-area modules have efficiencies ranging from 10% to 15%.

The highest solar cell efficiencies to date have been achieved in multijunction cells based on III–V compound semiconductors (Fig. 1).[4] These devices have multiple absorbing layers, each with different band gaps. Such a device scheme allows for higher theoretically attainable efficiencies. For devices with two, three, and "infinite" absorbing layers, the theoretical efficiency limits are 50%, 56%, and 85%, respectively.[5,6] The current record efficiency is 36.9% for a $Ga_{0.44}In_{0.56}P/Ga_{0.92}In_{0.08}As/Ge$ triple-junction research cell, measured under concentrated light. These devices are grown epitaxially on single crystal substrates using metal-organic chemical vapor

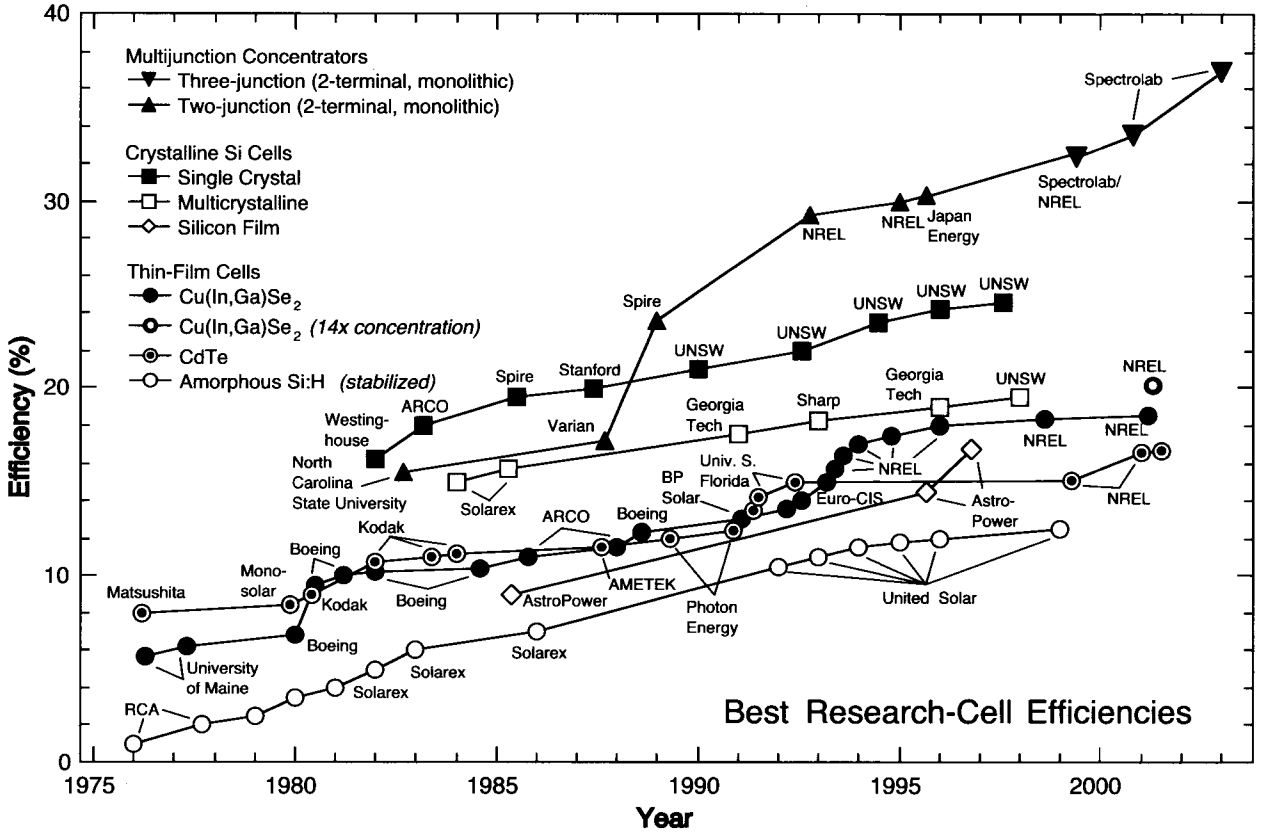


Fig. 1 Progress of photovoltaic device efficiencies (under AM1.5 simulation solar illumination) for a variety of inorganic technologies. (From Ref. [4].)

deposition or molecular beam epitaxy. Such highly sophisticated fabrication techniques are generally considered to be too expensive for large-scale commercialization and terrestrial photovoltaic energy production. However, these devices are used commercially in space-based applications, in which high efficiency is a necessity.

Over the last decade, there has been considerable effort in advancing thin-film, "second-generation" technologies that do not require the use of silicon wafer substrates and can therefore be manufactured at significantly reduced cost. Steady progress has been made in laboratory efficiencies (Fig. 1) for devices based on CdS/CdTe, $Cu(In,Ga)Se_2$ (CIGS), and amorphous Si. These devices are fabricated using techniques such as sputtering, physical vapor deposition, and hot-wire chemical vapor deposition. Multijunction cells based on amorphous-Si and amorphous-SiGe alloys have been the most successful second-generation technology to date because of their ability to be fabricated at relatively low cost and be integrated into electronics and building roofing material. They currently constitute most of the remaining 18% of the market not covered by single crystal silicon. Currently, there are about a dozen companies supplying photovoltaic modules based on thin-film technologies, with efficiencies ranging from 7% to 13.4%. It is anticipated that these second-generation technologies will rapidly grow in commercial importance as manufacturing cost and lifetime issues are improved.

Nonetheless, these existing technologies will be hard-pressed to meet the efficiency and cost goals for the long-term outlook. This has led to the concept first coined by Prof. Martin Green of the University of New South Wales of "third-generation" technologies which would be able to show dramatic improvements in either efficiency or cost or both, as shown in Fig. 2. To meet these ambitious goals will require the development of radically new devices that operate on fundamentally different principles than current technologies. A number of novel approaches have been suggested that may be capable, at least theoretically, of more closely approaching the thermodynamic limit of 86.8% for the conversion efficiency of sunlight to electricity.[7] These embryonic concepts include hot carrier solar cells, cells with multiple electron–hole pairs per photon, multiband quantum well and impurity solar cells, and thermophotovoltaic and thermophotonic devices. All of these concepts have embedded in them major technical challenges with respect to developing new devices, establishing a realistic efficiency potential for the devices, and demonstrating that they are inherently manufacturable.

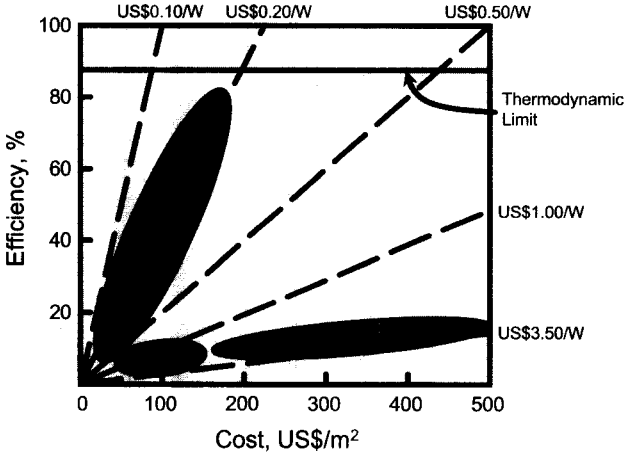


Fig. 2 Cost-efficiency matrix for first-, second-, and third-generation photovoltaic technologies. (From Ref. [7].) (*View this art in color at www.dekker.com.*)

Organic-Based Nanostructured Photovoltaics

The other approach to photovoltaic device development is to pursue technologies with moderate efficiency goals (<20%) but with significantly reduced fabrication costs (<$20/m^2). These are technologies that are potentially capable of the large-scale production necessary for a paradigm shift in global energy production. These would optimally be fabricated from solution-processible materials using low-temperature, nonvacuum techniques such as inkjet or screen printing. They would be compatible with low-cost, flexible substrates, thus allowing for high-throughput roll-to-roll processing as well as easy integration into existing building structures.

Organic molecular/polymeric semiconductors meet all of the above manufacturing process criteria. They are often soluble in common organic solvents and can be processed in a bench top laboratory setting. They have inherently low material costs and do not typically rely on the use of environmentally hazardous or elementally rare constituents. In addition, they have been proven to possess a wide range of optical, electronic, and mechanical properties that can be finely tuned through structural changes and molecular derivatization. In principle, they provide an enormous parameter space from which specific functionalities can be chosen to engineer novel, high-performance devices. These qualities make organic semiconductors strong candidates for a growing list of device applications. The area of flat-panel displays, in particular, is on the verge of a revolution because of advances in organic light-emitting diodes (OLEDs).[8,9] This technology is maturing rapidly, with first generation devices now on the market.

Using organic semiconductors in photovoltaic devices is now the subject of intense research. Organic molecules and π-conjugated polymers can have very high optical absorption coefficients ($>10^5$ cm^{-1}), making them useful as the active layer in the device. Additionally, charge carrier mobilities, while generally several orders of magnitude lower than in conventional semiconductors, are sufficiently high to effectively transport charge out of the device. A multitude of organic solar cell device designs is currently being investigated. All of these use some form of heterojunction between electron-donating and electron-accepting molecules to create free charge carriers from the bound electron–hole pairs (excitons) that are characteristic of organic semiconductors. To date, the reported power conversion efficiencies of these devices under simulated solar illumination are all less than 4%.

A great deal of basic research is needed to achieve significantly higher efficiencies in organic basic devices. Some of the key issues are:

- Exciton creation—The process of light absorption to create bound electron–hole pairs.
- Exciton diffusion—Diffusion of excitons to an interface where they can be dissociated.
- Electron transfer—Transfer of electrons to acceptor molecules.
- Charge transport—Transport of electrons and holes to the electrodes.

Many of these phenomena are determined by the morphologies of the molecular structures that occur at multiple length scales within the device. Morphologies on the length scale of a few nanometers, which affect exciton diffusion rates and charge carrier mobilities for instance, are determined by intermolecular forces and the kinetics of the growth/formation process of the material. Morphologies at longer length scales (tens to hundreds of nanometers), which affect the general direction of charge transport, are also determined by these factors, at least for an all-organic material. However, organic–inorganic composite systems allow the possibility of directly controlling the arrangement of the organic molecules at longer length scales. Inorganic materials, particularly semiconducting oxides, can be fabricated using a growing number of techniques to achieve specific three-dimensional structures. These structures can be used as a scaffolding onto which an organic material can be attached to form composite materials in which all the above photovoltaic device phenomena are optimized.

In this article, we give a brief description of the theory of charge transport in organic materials. We then provide an overview of the existing field of all-organic-based photovoltaic devices and go on to discuss current and hypothetical organic–inorganic composite devices. Finally, we examine a variety of novel semiconducting oxide structures that can be fabricated using low-temperature, solution- and electrochemical-based techniques.

CHARGE AND EXCITON TRANSPORT IN ORGANIC SEMICONDUCTORS

Organic semiconductors can be broadly categorized as being either small molecules (molecular weight less than a few thousand atomic mass units) or polymers (molecular weight between a few thousand and ~1 million amu). The distinction becomes most important in determining the processing that is required in making films and devices and for the subsequent morphologies that are obtained. However, the fundamental mechanisms that underlie light absorption, exciton diffusion, and charge carrier motion are essentially the same in the two classes of molecules. These properties are governed primarily by molecular orbitals that are built-up from π-electrons that are delocalized across the molecule. Examples of small molecules with extended π-electron delocalization and π-conjugated polymers commonly used in organic photovoltaic devices are depicted in Fig. 3. Highest occupied molecular orbitals (HOMOs) calculated for a sexithienyl oligomer, a fullerene, and a perylene diimide are shown in Fig. 4. Charge transfer and exciton diffusion processes are largely determined by the degree to which the orbitals of adjacent molecules and molecular subunits overlap.

Charge transfer between organic molecules can be viewed within the context of the semiclassical electron transfer theory of Marcus.[10] The rate of electron transfer between two molecules is given by

$$k_{ET} = \frac{4\pi^2}{\hbar}\frac{1}{\sqrt{4\pi k_B T}}t^2 \exp\left(-\frac{\lambda}{4k_B T}\right) \quad (1)$$

where $\hbar$ is Planck's constant, k_B is Boltzmann's constant, T is the temperature, t is the transfer integral describing the strength of interaction between the two molecules, and λ is the reorganization energy that describes the strength of the electron–phonon interaction.[11] Thus for fast charge transfer between molecules, a large interaction strength and small reorganization energy are required. Their relative orientation and the distance between them determine the degree of interaction between two molecules. For highly ordered, single- or polycrystalline molecular films of materials such as pentacenes or perylenes, this interaction strength can be sufficiently large for band transport to be observed.[11,12] However, for π-conjugated polymers and amorphous small-molecule materials, the large degree of disorder results in carrier bandwidths on the order of 10 meV.[13] which is far smaller than the ~0.1 eV required for band transport to

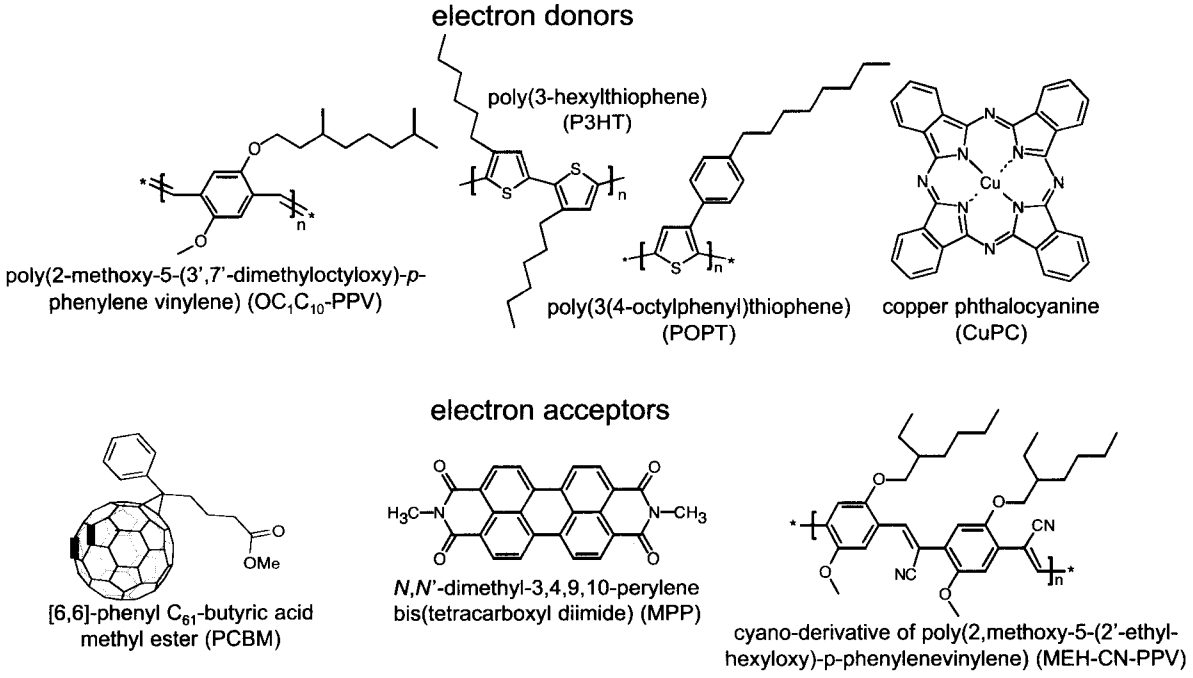


Fig. 3 Molecular structures of typical organic semiconductors that have been used in photovoltaic devices.

occur.[11] Instead, carrier transport in these materials occurs via polaron hopping between sites, where each site has a defined energy and geometric (positional+orientational) disorder parameter. In small-molecule materials, a "site" is considered to be an entire molecule. In conjugated polymers, however, the π-electrons, in reality, are not delocalized along the entire length of a polymer chain. Structural defects (kinks and bends) and chemical defects destroy the conjugation periodically; electronically, the polymer is a series of π-conjugated segments (typically 6–8 monomers long) that are bonded by nonconjugated sections. Fig. 5 illustrates carrier hopping between segments of a conjugated polymer.

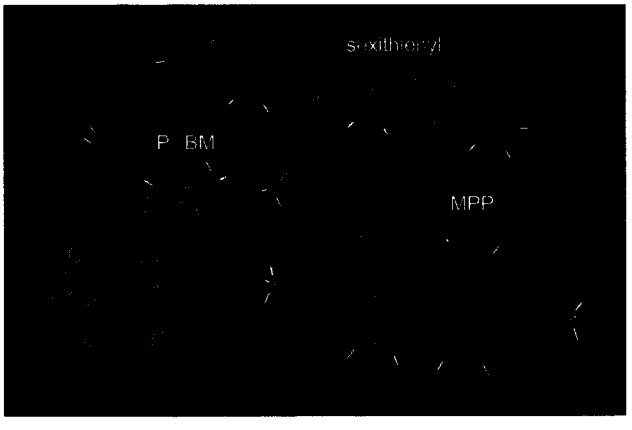


Fig. 4 Highest occupied molecular orbitals (from semiempirical calculations) of representative organic semiconductors with extended π-electron delocalization. (*View this art in color at www.dekker.com.*)

The charge carrier mobility for a broad range of organic semiconductors, both small molecule and polymer, has been found experimentally to follow Poole–Frenkel behavior, in which the mobility μ as a function of the electric field (E) and temperature is given by

$$\mu(E,T) \propto \mu_0(T)\exp\left(\gamma(T)\sqrt{E}\right) \tag{2}$$

where μ_0 is the zero-field mobility and $\gamma(T)$ is the coefficient of the stretched exponential that describes the temperature dependence of the field activation. This behavior has been successfully modeled with the "disorder formalism" originally developed by Bässler.[14] The model assumes hopping between sites with Gaussian distributions of energies and geometric disorder parameters. A more recent version of the model developed by Novikov et al.,[15] which takes into account spatial correlations in the site energies, results in

$$\mu = \mu_\infty \exp\left[-\left(\frac{3\hat{\sigma}_d}{5}\right)^2 + C_0(\hat{\sigma}_d^{3/2} - \Gamma)\sqrt{\frac{eaE}{\hat{\sigma}_d}}\right] \tag{3}$$

where μ_∞ is the mobility in the limit $T\to\infty$, $\hat{\sigma}_d$ is the width of the Gaussian distribution of site energies divided by k_BT, C_0 is an empirical constant, Γ describes the

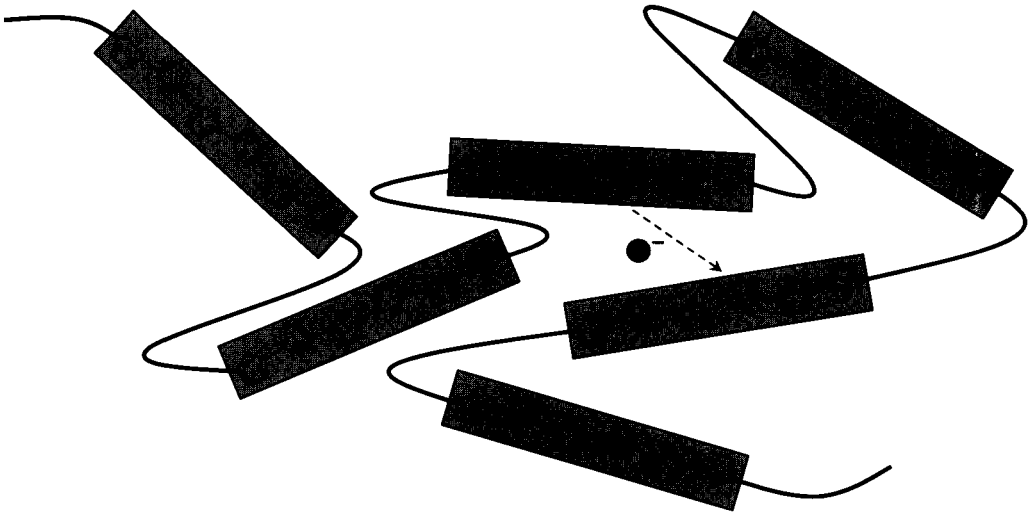


Fig. 5 Conceptual diagram of the mechanism of charge transport by hopping in a π-conjugated polymer. The rectangular sections represent regions of the polymer that are fully conjugated. Between these conjugated segments are regions of polymer that are bent or twisted, or that have chemical defects that destroy the conjugation. (*View this art in color at www.dekker.com.*)

geometric disorder, and a is the intersite spacing. Using typical values of ~0.1 eV and 2 for $\hat{\sigma}_d$ and Γ, respectively, it becomes obvious that the mobility in these materials increases as the degree of disorder is decreased. Typical charge carrier mobilities in conjugated polymers that have been reported in the literature are in the range of 10^{-6}–10^{-3} cm^2 V^{-1} sec^{-1}.[16]

Exciton motion/diffusion in organic semiconductors is determined by the same intermolecular interaction considerations as carrier transport, and thus exciton diffusion rates can also be described by hopping models involving Gaussian distributions of site energies. This has been demonstrated by thermally stimulated photoluminescence experiments on conjugated polymer films, which have been successfully modeled using double-peak Gaussian distributions to account for regions of aggregated polymer.[17] Calculations have shown that the electronic coupling between conjugated segments that are sequentially adjacent along the backbone of the polymer (that is, physically bonded but electronically isolated by a break in the conjugation) is approximately 1 order of magnitude less than the coupling between conjugated segments on different polymer chains (or on the same chain, but far away sequentially), which are close together and aligned parallel to each other.[18] This has lead to the conclusion that interchain exciton hopping is a faster process than intrachain exciton hopping.[19] Analogous to the limitations on carrier transport, poor interaction strengths between adjacent segments typically lead to short exciton diffusion lengths (the average distance the exciton moves before decaying to the ground state or becoming dissociated at a defect or interface) of the order of 6–10 nm,[20] although much large exciton diffusion lengths have been reported for highly ordered materials.[21]

The main conclusion to be drawn from this section is that molecular ordering is a critical ingredient to obtain high carrier mobilities and large exciton diffusion lengths in organic semiconductors. This point is illustrated in the polymer poly(3-hexylthiophene) (P3HT). This material (when synthesized in its regioregular, head-to-tail form) has been shown to possess a large degree of molecular ordering because of efficient interdigitation of the alkyl chains of adjacent polymer chains, leading to strong π-stacking of the thiophene backbones (Fig. 6). As a result, this material has shown the highest hole mobility of a conjugated polymer to date (0.1 cm^2 V^{-1} sec^{-1}), when measured in a field effect transistor geometry.[22] The effect of this molecular ordering on photovoltaic device performance will be discussed in the next section.

ORGANIC PHOTOVOLTAIC DEVICES: THE PLASTIC SOLAR CELL

Photoexcitations in organic semiconductors are inherently different than in conventional inorganic semiconductors. Whereas light absorption in an inorganic semiconductor typically leads to the immediate production of a free electron–hole pair, light absorption in an organic semiconductor results in the formation of an exciton (bound electron–hole pairs). In order for a free carrier to be

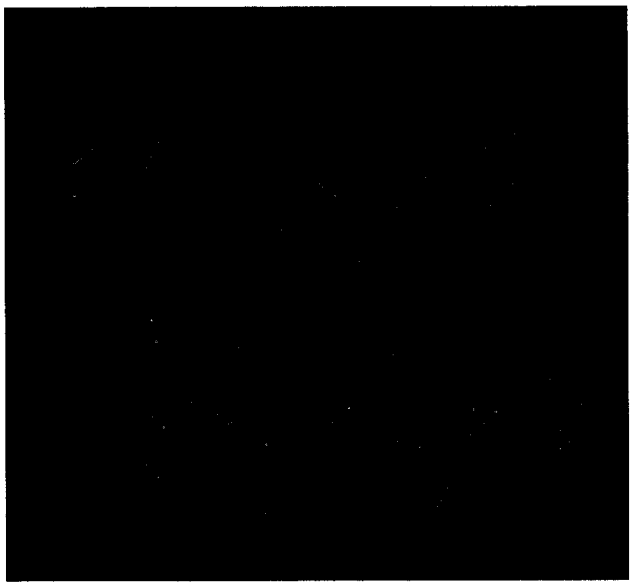


Fig. 6 Conceptual diagram of interchain ordering in a P3HT aggregate (cyan=carbon, yellow=sulfur, hydrogen atoms are not shown). (*View this art in color at www.dekker.com.*)

generated, this exciton must be dissociated. This can happen in the presence of very high electric fields, at a defect site in the material, or at the interface between two materials that have a sufficient mismatch of their energetic levels (band offset). In the case of dissociation at the interface between two materials, an exciton created in the material with lower (higher) electron affinity diffuses to the interface, and electron (hole) transfer occurs to the material with higher (lower) electron affinity. The higher electron affinity material is referred to as the acceptor, while the lower electron affinity material is the donor. Thus one can fabricate a photovoltaic device with the structure positive electrode/donor/acceptor/negative electrode. Tang[23] using a copper phthalocyanine layer as the donor and a perylene derivative as the acceptor first did this. This device had a power conversion efficiency of about 1% under simulated solar illumination.

There are two possible driving forces in the operation of a photovoltaic device: the electrical potential gradient (the electric field) and the chemical potential gradient.[24] In a bilayer donor–acceptor device, a chemical potential gradient is established that preferentially drives electrons away from the interface toward the negative electrode and holes away from the interface toward the positive electrode. An electric field in the device, such as one established by a difference in the work functions of the positive and negative electrodes, can assist in driving charges toward the electrodes, but may not be critical to the operation of the device. The important criteria for high efficiency in such a device are large exciton diffusion lengths in both the donor and acceptor materials, fast electron transfer and slow recombination rates at the interface, and high charge carrier mobilities in both layers. Several molecular donor–acceptor pairs have been investigated for use in bilayer photovoltaic devices, including phthalocyanine–perylene,[23,25] phthalocyanine–fullerene,[26,27] polymer–polymer,[28] polymer–perylene,[29] and polymer–fullerene.[30] As a result of the short exciton diffusion lengths in these materials, only excitons that are generated within approximately 10 nm of the interface contribute to the photocurrent. This presents a serious limitation on the possible power conversion efficiency of these devices.

To overcome this problem, researchers have developed devices in which the donor and acceptor molecules are blended together in an intimate mixture to produce a "bulk heterojunction" solar cell.[31] The donor and acceptor molecules form two "interpenetrating networks" that transport holes and electrons, respectively. In this device, an exciton generated anywhere in the bulk has only a few nanometers to diffuse before encountering an acceptor interface. After dissociation of the exciton into a free electron and a free hole, the charge carriers are transported through the bulk of the film (along their respective transport networks) to the electrodes. Directionality of the charge transport is provided by a built-in electric field in the device, which originates from a difference in effective work functions of the electrodes, as well as possibly a chemical potential gradient, which originates from "selective" electrodes that preferentially block either holes or electrons.[32]

Bulk heterojunction solar cells based on blends of a conjugated polymer with a fullerene (Fig. 7) have been among the most successful to date.[33] Photoinduced electron transfer from a conjugated polymer to a fullerene was demonstrated to occur on an ultrafast timescale ($\sim$45 fsec) and to produce a long-lived charge-separated state, with recombination to the ground state following a power law decay ranging from 300 nsec to 1 msec.[34,35] Thus excitons generated in the conjugated polymer, which have lifetimes of several hundred picoseconds, are quickly dissociated into free carriers when they encounter a fullerene interface. These free carriers will have sufficient time to be transported to the electrodes before recombining provided that the active layer of the device is not too thick. This produces an inherent tradeoff in the device design. As the active layer is made thicker, more light is absorbed and more excitons/charge carriers are created, but the carriers spend longer time in the device and thus have a higher probability of recombining to the ground state before they can contribute to the photocurrent. This means that high carrier mobility (for both holes and

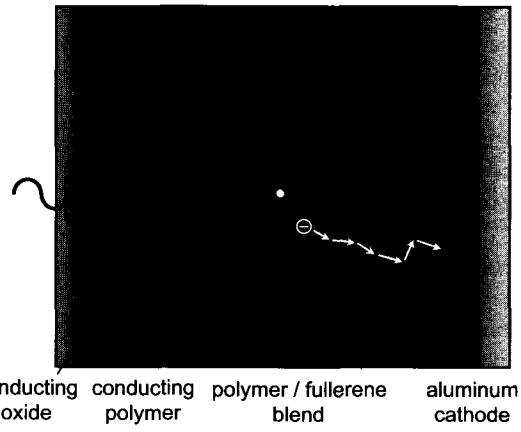


Fig. 7 Schematic diagram of a conjugated polymer–fulllerene bulk heterojunction photovoltaic device. (*View this art in color at www.dekker.com.*)

electrons) is required in the elements of the blend to allow the active layer to be thick enough to have significant optical density.

As discussed above, morphology is an important factor in determining a material's carrier mobility. Thus a critical area for development of high efficiency devices is to develop approaches that lead to improved morphologies. This has been demonstrated in conjugated polymer–fullerene bulk heterojunction solar cells in which changing the casting solvent for the active layer film can lead to a large increase in device photocurrent and efficiency.[36] Devices consisting of a blend of the polymer poly(2-methoxy-5-(3′,7′-dimethyloctyloxy)-*p*-phenylene vinylene) (OC_1C_{10}-PPV) with the fullerene [6,6]-phenyl C_{61}-butyric acid methyl ester (PCBM) (Fig. 3) showed nearly a factor of 3 enhancement in efficiency when cast from chlorobenzene as opposed to toluene. It was subsequently determined that the hole mobility in OC_1C_{10}-PPV is approximately 1 order of magnitude higher when cast from the chlorobenzene. From light-scattering measurements performed over a range of polymer concentrations, it was shown that the polymer does not truly dissolve in chlorobenzene, but instead forms a suspension.[37] A complete explanation for why this leads to higher carrier mobilities in the thin-film material has not been fully elucidated, but it has been postulated that the polymer chains collapse together into an aggregate that promotes π-electron sharing (π-stacking) between chains. Such solvent effects on carrier mobilities have also been demonstrated in the operation of OLEDs.[19] Electron mobilities in the bulk heterojunction OC_1C_{10}-PPV–PCBM blend have also been shown to be high (2×10^{-3} cm^2 V^{-1} s^{-1}) when the film is cast from chlorobenzene, although, in this case, the explanation apparently is that chlorobenzene is a very good solvent for PCBM and reduces the amount of phase segregation that normally leads to large barriers to electron transport in the fullerene network.[38] Additional evidence for the effects of morphology on mobility and device performance has been reported in bulk heterojunction solar cells made from the conjugated polymer P3HT (Fig. 3). It was demonstrated that post-production annealing of photovoltaic devices made from P3HT–PCBM blends results in a dramatic increase in device efficiency, from 0.4% to 3.5% as measured under white light illumination. This efficiency enhancement presumably arises from an increase in the ordering of the P3HT chains upon annealing.[39]

The general viability of the application of donor–acceptor systems as employed in bulk heterojunction organic solar cells has been demonstrated by a number of other recent results, including polymer–polymer[40] and polymer–quantum rod devices.[41] These all have the same criteria for high efficiency as the polymer–fullerene device, namely, small distances for exciton diffusion to an interface, exciton dissociation that is fast, charge recombination that is slow, and large charge carrier mobilities. Reported efficiencies of organic solar cells to date have not exceeded 4% for several reasons. First, the amount of light that is absorbed by the devices is typically too small because 1) the optical band gap of the light-absorbing component is usually too large (>1.4 eV), resulting in poor overlap with the solar spectrum, and 2) the optical density is too small because the active layer of the device must be kept thin to limit charge carrier recombination. Second, the photovoltage that is generated is typically too low. This is limited, to first order, by the effective electronic band gap of the active layer, which is given by the difference between the HOMO level of the donor and the lowest unoccupied orbital (LUMO) level of the acceptor. Last, the fill factor, which is a measure of the quality of the diode, is too small. A necessary solution for many of these problems will be the development of new polymers and organic molecules. Concurrently, devising new ways to control the morphologies of the light-absorbing and -transporting materials to optimize the pathways for exciton/charge transport will be required. The organic–inorganic composite devices discussed in the next section provide a promising route to controlling morphologies on the nanoscale.

ORGANIC–INORGANIC COMPOSITE PHOTOVOLTAIC DEVICE CONCEPTS

The bulk heterojunction solar cell is one approach to addressing the problem of short exciton diffusion lengths in organic semiconductors. However, the mixing of donor

and acceptor species into an intimate blend introduces the problem of bulk recombination of carriers to the ground state, as previously discussed. Additionally, islands, defects, and cul-de-sacs in the transport networks that result in incomplete percolation paths may limit charge transport of one or both carriers in a bulk heterojunction cell. One promising route to high-efficiency solar cells that overcome these problems are devices that use a nanostructured, semiconducting metal oxide (such as TiO_2 SnO_2, or ZnO) as the electron-accepting and -transporting network. Such materials can be easily grown using low-temperature, solution-based routes to produce a variety of nanostructures, as is discussed in the next section. The topology of these nanostructures can be optimized to produce a continuous network that guarantees a percolation path for electrons to the electrode. This then opens up the possibility of organic–inorganic composite devices in which an organic semiconductor is deposited onto the surface of the oxide or used to fill the empty volume in a porous structure. In this way, the inorganic network can be used to induce a specifically tailored morphological structure in the organic material and result in an overall device structure that has a large optical path length but short carrier-to-electrode path lengths.

The first example of a solar cell made using a porous oxide as an electron-transporting network was the solution-based, dye-sensitized solar cell.[42] In this device, a monomolecular layer of dye (typically a ruthenium complex) is chemisorbed through carboxylic acid binding groups onto the surface of a high surface area, mesoporous TiO_2 film. A liquid electrolyte (typically with an iodine/triiodide redox couple in an organic solvent) fills the pore spaces and makes contact to a platinum-coated counter-electrode. In the operation of the device, the dye molecules absorb light and inject electrons into the TiO_2. The oxidized dye molecules are subsequently reduced by a reaction with the redox species. Transport of the charges out of the device occurs via diffusion of electrons through the TiO_2 and diffusion/ionic conduction of the redox species to the counterelectrode. Such devices have achieved 10% power conversion efficiency under solar illumination; however, the use of a liquid electrolyte makes them very difficult to implement for large-scale production. Recent progress has been made in replacing the liquid electrolyte with organic semiconductors, yielding devices with efficiencies reaching 3.2%.[43,44] Also, preliminary investigations into sensitizing a mesoporous oxide with quantum dots have been performed,[45] opening up the possibility of harnessing some of the "third-generation" photovoltaic mechanisms for achieving higher efficiencies such as hot carrier production or impact ionization.[46]

More recently, researchers have investigated solar cells based on conjugated polymer–semiconducting oxide composites. Such devices consist of a conjugated polymer embedded into a mesoporous metal oxide, which acts as an electron acceptor. Several studies have shown that metal oxide semiconductors can act as efficient electron acceptors for conjugated polymers.[47,48] The operation of these devices differs from that of the dye-sensitized cell in that light is absorbed in regions of the conjugated polymer potentially far from an oxide interface. Thus excitons must diffuse to an interface before electron transfer to the oxide can take place. Also, electric fields that contribute to the driving force for the current may be present in the device, which is not the case in the dye-sensitized cell because the ionic liquid shields all electric fields.[24] Several possible morphologies for such a composite device can be envisioned. As shown in Fig. 8, the polymer can be embedded in an isotropic, mesoporous oxide or in a "forest" of oxide nanorods that are vertically oriented to the substrate. To facilitate a direct path for charges to be transported to the electrode, the vertically oriented columns may be a better geometry. Aside from the goal of increased power conversion efficiency, these composite devices may result in increased morphological stability

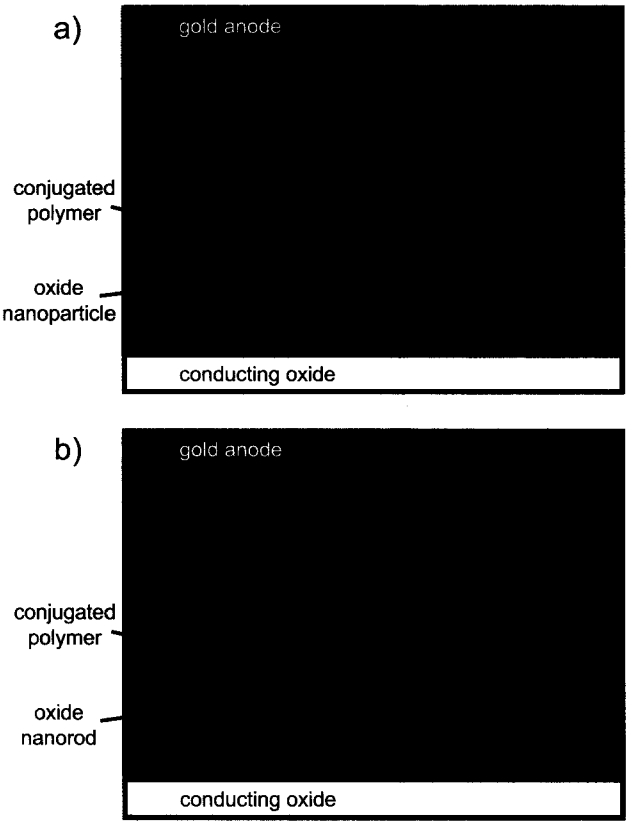


Fig. 8 Schematic diagram of conjugated polymer–semiconducting oxide composite photovoltaic devices with a) an isotropic, mesoporous oxide and b) a forest of oxide nanorods. (*View this art in color at www.dekker.com.*)

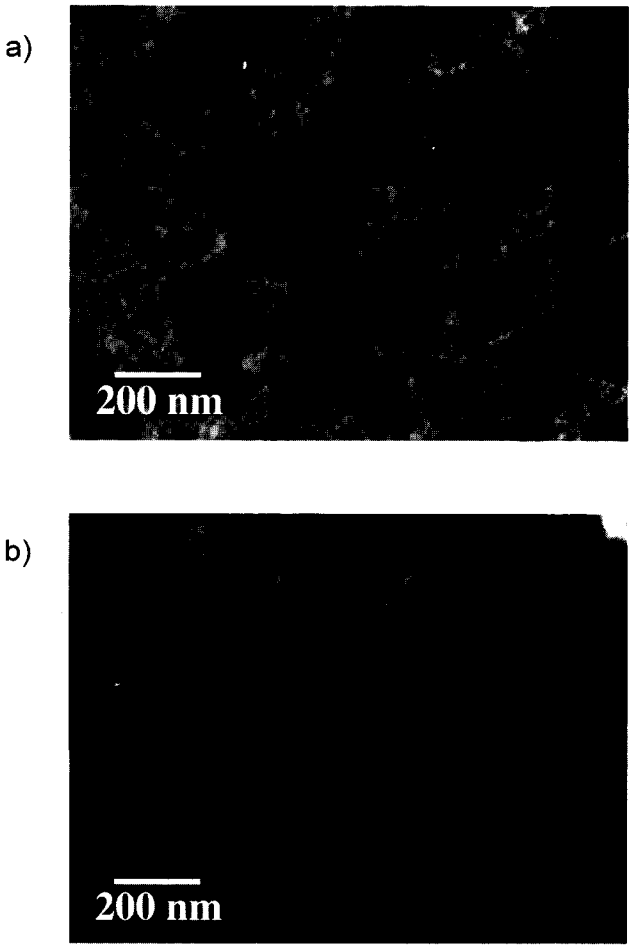


Fig. 9 Scanning electron microscopic images of the top surfaces of a) a porous SnO_2 film with a thickness of 300 nm and b) a SnO_2/conjugated polymer (OC_1C_{10}-PPV) composite film. The porous SnO_2 film was fabricated by sintering a solution-deposited SnO_2 nanoparticle/polystyrene sphere mixture. The composite structure was fabricated by solution intercalation of the polymer into the porous SnO_2 structure.

and robustness compared with pure polymer films, resulting in enhanced device lifetimes.

Fabrication of organic–inorganic composite devices has been performed via several techniques. Fig. 9a shows a mesoporous SnO_2 structure that was fabricated by spin-coating an aqueous solution of SnO_2 nanoparticles (diameter$\sim$9 nm) mixed with polystyrene spheres (diameter$\sim$100 nm). Upon heating the film to 450°C in an oxygenated atmosphere, the SnO_2 nanoparticles are sintered together to form a continuous oxide network. The polystyrene spheres are burned out of the film, leaving behind voids on the size scale of 100 nm. Fig. 9b shows a composite structure fabricated by soaking the above film in a solution of the conjugated polymer OC_1C_{10}-PPV for 8 hr. The SEM image shows efficient intercalation of the polymer into the porous SnO_2 structure. From optical measurements, the polymer was estimated to fill 75% of the free volume of the oxide. Several thermodynamic factors contribute to the degree of intercalation of a polymer into a porous medium.[49] Osmotic pressure drives the polymer chains into the pores. However, the entropic force is in the other direction because of the reduced number of configurational states available to the confined polymer. The force that can tip the scales in either direction is the interaction energy between the polymer and oxide surface. A strong attractive force between the polymer and the oxide can overcome the entropic drive and result in efficient intercalation into the porous volume. Whether the oxide–polymer interaction is attractive or repulsive depends on a wide range of intermolecular forces, including the hydrophobic/hydrophilic and polar character of both materials, as well as the solvent in which the polymer is dissolved.[50] The hydrophobic/hydrophilic nature of the oxide surface can be substantially altered by cleaning processes or chemical derivatization to optimize the conditions for polymer intercalation. In a system in which there is a strong oxide–polymer attraction leading to efficient intercalation, the intercalated polymer can become essentially insoluble in solvents that would normally dissolve it, as is the case for the composite structure shown in Fig. 9b.

Other techniques that have been demonstrated include a self-assembly approach to fabrication of mesoporous TiO_2 using a block copolymer as a structure-directing agent.[51] Fig. 10 shows a highly ordered, mesoporous TiO_2 structure fabricated via this technique. Intercalation

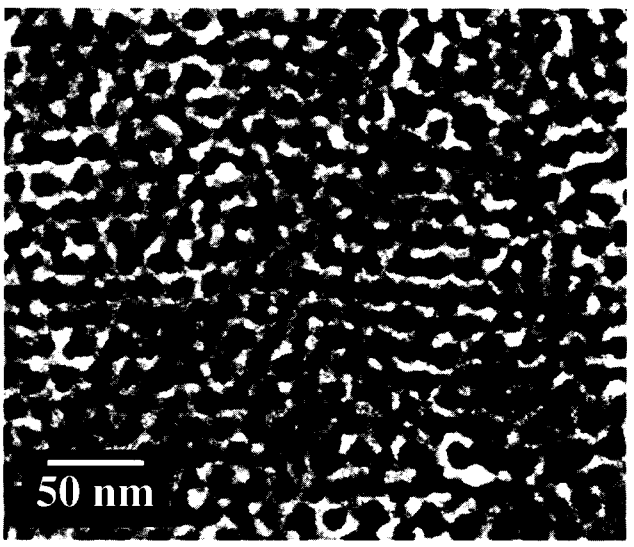


Fig. 10 Scanning electron microscopic image of an ordered, porous TiO_2 film fabricated using a self-organizing block copolymer as a templating agent. (From Prof. M.D. McGehee, Dept. of Materials Science and Engineering, Stanford University.)

of this structure with the conjugated polymer P3HT resulted in a solar cell with approximately 0.5% power conversion efficiency under solar illumination. Also, a technique in which a $Ti(OC_3H_7)_4$ precursor was blended with a conjugated polymer and subsequently annealed to form a composite film with 100% filling has been demonstrated.[48] In the next section, we examine a variety of novel, low-cost techniques for fabricating metal oxide nanostructures that could potentially be used in an organic–inorganic composite solar cell.

NOVEL FABRICATION TECHNIQUES FOR METAL OXIDE NANOSTRUCTURES

Over the past few years, the synthesis of semiconducting nanoparticles and nanorods has advanced substantially. There is considerable work demonstrating the synthesis of nearly monodisperse particles with sizes from a few nanometers to a few hundred nanometers. This allows tremendous flexibility in controlling the material band structure. There has been a remarkable diversity of approaches to the synthesis of nanoparticles of various structures and of nanostructured surfaces. These include solution precipitation, sol–gel approaches, self-organization, and a range of physical vapor deposition (PVD)-based approaches including pulsed laser deposition, chemical vapor deposition, and sputtering.[52–61] One of the key emerging areas for controlling particle synthesis and incorporation into composite structures is the use of appropriate chelating ligands to promote self-organization.[53,62–70] There is quite a remarkable diversity of approaches being employed to both produce and tailor nanoparticulate materials. This provides an unprecedented degree of flexibility in the development of new materials. Interestingly, however, the vast majority of the work has

Table 2 Synthetic routes for the formation of nanofibers/nanotubes from a variety of inorganic semiconductors

Material	Synthetic method	Dimensions	Reference
H-TiO_2 nanotubes	Hydrothermal 10 M NaOH	5 nm radius, 150 nm length	[82,83]
TiO_2 nanofibers	Electrospinning of ethanol solution of poly(vinyl pyrrolidone)/titanium tetraisopropoxide	10–100 nm radius, up to several centimeters length	[84]
InAs nanofibers	$2InCl_3+6KBH_4+2As \rightarrow 2InAs+6KCl+6\ BH_3+3H_2$	15–100 nm radius, 150–1000 nm length	[85]
GaP nanorods	Benzene reflux with $GaCl_3$, Na, and P	20–50 nm radius, 200–500 nm length	[86]
CdE (E=S, Se, Te) nanorods	Direct reaction of the elements in organic solvents	10–40 nm radius, microns length	[87]
CuInE (E=S, Se) nanorods	Elements in organic solvents	10–40 nm radius, hundreds of nanometers length	[88]
ZnO, ZnS nanocables and nanotubes	Evaporation of the oxide followed by sulfiding	5–20 nm radius, hundreds of nanometers length	[89]
ZnO nanorods	Thermal decomposition of Zn^{II} amino complex in aqueous solution	5–10 nm, 50–100 nm, and 0.5–1.0 μm radius, up to 10 μm length	[90,91]
ZnO nanobridges and nanonails	Thermal vapor transport and condensation method	25–100 nm radius, hundreds of nanometers to 2 μm length	[92]
GaP nanorods	Catalytic decomposition of Ga_2O, P, and carbon nanotubes	50 nm radius, microns length	[93]
GaO nanoribbons	Direct reaction of GaN with oxygen	10–60 nm thick, hundreds of nanometers length	[94]
MnO_2 nanowires	Hydrothermal reaction of $MnSO_4$–H_2O with water in presence of ammonium persulfate	5–20 nm radius, microns length	[95]
$CdWO_4$ nanorods	Hydrothermal preparation	50 nm radius, 200 nm length	[96]
$BaWO_4$ nanorods	Reverse micelle synthesis	20–50 nm radius, hundreds of nanometers length	[97,98]
Vanadium oxide nanotubes	Hydrothermal synthesis with ammonium metavanadate and amines or quaternary ammonium salts	20–100 nm radius, 100–1000 nm length	[99]
MoS_2, WS_2, and NbS_2 nanotubes	High temperature synthesis of Mo with Se	100 nm radius, microns length	[100–103]

been on a very few materials such as CdS, CdSe, silica, and ZnO.[62,71–80] This work served well to develop the basic tools for the synthesis and capping of the nanoparticles for self-assembly or incorporation into composites. Recent work has also developed an array of tools for the synthesis of core/shell nanoparticles where the surface of the particle is a different material than the interior.[56,73,81] Overall, the development of these synthetic techniques provides the opportunity to pick a core material based on optical, electronic, and chemical properties and then cap it with another electronic material or to derivatize it with appropriate ligands to promote incorporation into a composite, or to lead self-assembly, or to control exciton or charge transfer properties. This presents an unprecedented ability to tailor the behavior of a composite material and has significant ramifications for a wide range of applications including solar cells, catalysis, and sensors.

One of the newer developments in this area has been the development of synthetic approaches to nanorods or tubes with high aspect ratios. This leads to the possibility of anisotropic three-dimensional structures. The rods can be organized very differently than spherical particles and can be selectively attached or derivatized to control incorporation into a composite. As photoactive elements, nanomaterials with high aspect ratios can be used for transport of carriers and selective transport along the fiber or tube. For tubes, it may also be possible to delineate between the interior and exterior of the tube. The formation of nanorods and or tubes has been observed for a wide variety of systems, some of which are indicated in Table 2. While this list is not exhaustive by any means, it should serve to illustrate the extensive range of possible materials available with a wide variety of surface chemistries, optical, and electronic properties. Key is that simple hydrothermal or gas phase synthesis routes, which are very flexible allowing for derivatization of the materials and are very scalable, can make many of the materials. Figs. 11 and 12 show examples of TiO_2 nanofibers and ZnO microcrystallites grown in aqueous solutions.

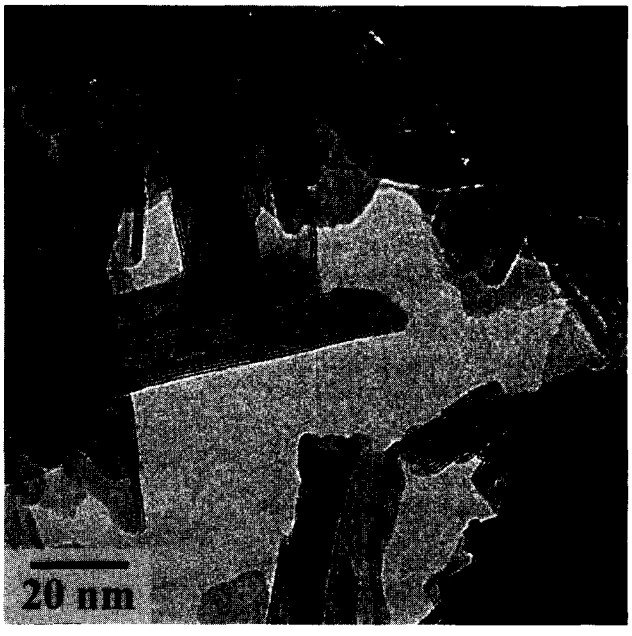


Fig. 11 Scanning electron microscopic image of TiO_2 nanofibers grown from an aqueous solution at 150°C.

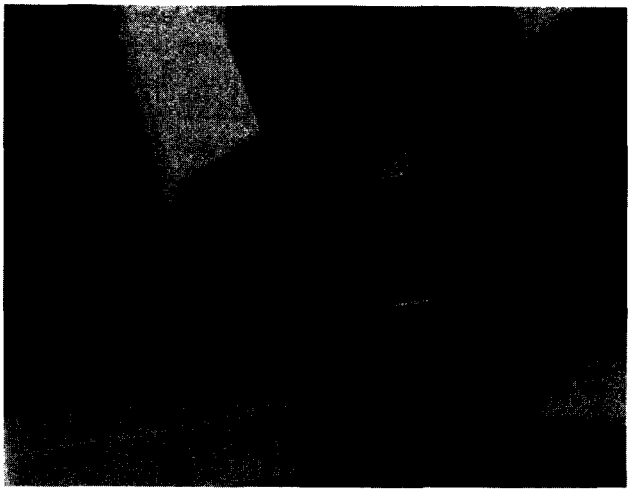


Fig. 12 Scanning electron microscopic image of ZnO microcrystallites grown from an aqueous solution at 95°C.

There is also a very embryonic literature on the photochemistry of nanorods/tubes indicating that very facile charge transport can occur with appropriate surface derivatization (linkers) and produce useful chemistry such as polymerization, etc.[104–108] These results support the ability to apply the tunability of the semiconducting nanostructures through quantum confinement, coupled with rapid electron/exciton transfer with appropriate incorporation of the materials into the composite structure. These initial results point to the need for extensive work to understand the nature of excitonic and charge transfer in inorganic nanorods and nanotubes as well as the specifics of how to couple the nanomaterial to the outside environment such as a polymer or catalyst entity. However, there is every reason to believe at present that the diversity of materials and ligands will allow tailoring of the energetics and kinetics to produce specific redox behavior.

CONCLUSION

Organic semiconductors are a promising class of materials for the fabrication of low-cost, large-scale production

solar cells. Current state-of-the-art organic solar cells have achieved power conversion efficiencies exceeding 3%. This is an encouraging value, but a substantial increase in performance is required for these devices to be of practical use. To achieve higher efficiencies will require controlling the morphologies of the donor and acceptor materials to simultaneously optimize the exciton and charge transport pathways while allowing for an optically thick device. Recent advances in solution-based synthesis of metal oxide and other inorganic semiconductor nanostructures bring about the possibility of low-cost organic–inorganic composite devices with optimized morphologies. Many challenges exist in this area however. Further techniques for fabricating composite structures must be identified and refined, particularly with respect to efficiently intercalating polymers into nanostructured/porous media. Also, correct alignment of the charge transport pathways to the electrodes is necessary for efficient device operation, and thus the orientation of the nanostructures with respect to the substrate must be precisely controlled. Perhaps the most vital, yet least understood, topic in this area is the organic–inorganic interface. Optimized device structures will require good mechanical contact and strong adhesion between the two components. Electronic properties of the interface are also crucial. Exciton quenching at the interface must be fast and efficient, and the relative electronic levels of the two materials must be optimized to yield the largest possible chemical potential difference across the interface for photogenerated charges. All these require a complete understanding of the chemistry and charge transfer reactions that occur at the interface. Meeting these challenges will be necessary to realize the full potential of organic semiconductor and organic–inorganic composite devices.

ACKNOWLEDGMENTS

The authors thank Alex Miedaner for synthesis of TiO_2 and ZnO nanostructures and Bobby To for production of SEM images. The authors gratefully acknowledge funding from the U.S. Department of Energy, the National Center for Photovoltaics, and from an NREL Director's Discretionary Research and Development grant under the Department of Energy contract DE-AC36-19GO10337.

REFERENCES

1. *Annual Energy Review 2001*; Energy Information Administration. http://www.eia.doe.gov/emeu/aer/elect.html (accessed February 2003).
2. Shockley, W.; Queisser, H.J. Detailed balance limit of efficiency of p–n junction solar cells. J. Appl. Phys. **1961**, *32* (3), 510–519.
3. Green, M.A.; Emery, K.; King, D.L.; Sanekazu, I.; Wilhelm, W. Solar cell efficiency tables (version 19). Prog. Photovolt.: Res. Appl. **2002**, *10*, 55–61.
4. Kazmerski, L.L. 21st Century Photovoltaics R&D: An Expedition Toward the Non-conventional. In *Photovoltaics for the 21st Century*; McConnell, R.D., Kapur, V.K., Eds.; The Electrochemical Society: Washington, DC, 2001; 20–29.
5. Sze, S.M. *Physics of Semiconductor Devices,* 2nd Ed.; John Wiley & Sons: New York, 1981.
6. Marti, A.; Araujo, G.L. Limiting efficiencies for photovoltaic energy conversion in multigap systems. Sol. Energy Mater. Sol. Cells **1996**, *32*, 203–222.
7. Green, M.A. Third Generation Photovoltaics: Concepts for High Efficiency at Low Cost. In *Photovoltaics for the 21st Century*; Kapur, V.K., Ed.; The Electrochemical Society: Washington, DC, 2001; 3–19.
8. Tang, C.W.; VanSlyke, S.A. Organic electroluminescent diodes. Appl. Phys. Lett. **1987**, *51* (12), 913–915.
9. Friend, R.H.; Gymer, R.W.; Holmes, A.B.; Burroughes, J.H.; Marks, R.N.; Taliani, C.; Bradley, D.D.C.; Dos Santos, D.A.; Bredas, J.L.; Logdlund, M.; Salaneck, W.R. Electroluminescence in conjugated polymers. Nature **1999**, *397* (6715), 121–128.
10. Marcus, R.A. Electron transfer reactions in chemistry. Theory and experiment. Rev. Mod. Phys. **1993**, *65* (3), 599–610.
11. Bredas, J.L.; Calbert, J.P.; da Silva Filho, D.A.; Cornil, J. Organic semiconductors: A theoretical characterization of the basic parameters governing charge transport. Proc. Natl. Acad. Sci. **2002**, *99* (9), 5804–5809.
12. Karl, N. Charge carrier transport in organic semiconductors. Synth. Met. **2003**, *133–134* 649–657.
13. Silinsh, E.A. *Organic Molecular Crystals. Their Electronic States*; Springer-Verlag: Berlin, 1980; Vol. 16.
14. Bässler, H. Charge transport in disordered organic photoconductors. Phys. Status Solidi, B **1993**, *175*, 15–56.
15. Novikov, S.V.; Dunlap, D.H.; Kenkre, V.M.; Parris, P.E.; Vannikov, A.V. Essential role of correlations in governing charge transport in disordered organic materials. Phys. Rev. Lett. **1998**, *81* (20), 4472–4475.
16. Martens, H.C.F.; Blom, P.W.M.; Schoo, H.F.M.

Comparative study of hole transport in poly(*p*-phenylene vinylene) derivatives. Phys. Rev., B **2000**, *61* (11), 7489–7493.
17. Kadashchuk, A.; Skryshevski, Y.; Piryatinshki, Y.; Vakhnin, A.; Emelianova, E.V.; V.I., A.; Bässler, H.; Shinar, J. Thermally stimulated photoluminescence in poly(2,5-dioctoxy *p*-phenylene vinylene). J. Appl. Phys. **2002**, *91* (8).
18. Beljonne, D.; Pourtois, G.; Cilva, C.; Hennebicq, E.; Herz, L.M.; Friend, R.H.; Scholes, G.D.; Setayesh, S.; Müllen, K.; Bredas, J.L. Interchain vs. intrachain energy transfer in acceptor-capped conjugated polymers. Proc. Natl. Acad. Sci. **2002**, *99* (17), 10982–10987.
19. Nguyen, T.-Q.; Martini, I.B.; Liu, J.; Schwartz, B.J. Controlling interchain interactions in conjugated polymers: The effects of chain morphology on exciton–exciton annihilation and aggregation in MEH-PPV films. J. Phys. Chem., B **2000**, *104*, 237–255.
20. Theander, M.; Yartsev, A.; Zigmantas, D.; Sundström, V.; Mammo, W.; Andersson, M.R.; Inganäs, O. Photoluminescence quenching at a polythiophene/C60 heterojunction. Phys. Rev., B **2000**, *61* (19), 12957–12963.
21. Gregg, B.A.; Sprague, J.; Peterson, M.W. Long-range singlet energy transfer in perylene bis(phenethylimide) films. J. Phys. Chem., B **1997**, *101* (27), 5362–5369.
22. Sirringhaus, H.; Brown, P.J.; Friend, R.H.; Nielsen, M.M.; Bechgaard, K.; Langeveld-Voss, B.M.W.; Spiering, A.J.H.; Janssen, R.A.J.; Meijer, W.W.; Herwig, P.; de Leeuw, D.M. Two-dimensional charge transport in self-organized, high-mobility conjugated polymers. Nature **1999**, *401*, 685–688.
23. Tang, C.W. Two-layer organic photovoltaic cell. Appl. Phys. Lett. **1986**, *48* (2), 183–185.
24. Gregg, B.A.; Hanna, M.C. Comparing organic to inorganic photovoltaic cells: Theory, experiment, and simulation. J. Appl. Phys. **2003**, *93* (6), 3605–3614.
25. Wohrle, D.; Meissner, D. Organic solar-cells. Adv. Mater. **1991**, *3* (3), 129–138.
26. Pannemann, C.; Dyakonov, V.; Parisi, J.; Hild, O.; Wohrle, D. Electrical characterisation of phthalocyanine–fullerene photovoltaic devices. Synth. Met. **2001**, *121* (1–3), 1585–1586.
27. Yakimov, A.; Forrest, S.R. High photovoltage multiple-heterojunction organic solar cells incorporating interfacial metallic nanoclusters. Appl. Phys. Lett. **2002**, *80* (9), 1667–1669.
28. Granström, M.; Petritsch, K.; Arias, A.C.; Lux, A.; Andersson, M.R.; Friend, R.H. Laminated fabrication of polymeric photovoltaic diodes. Nature **1998**, *395*, 257–260.
29. Breeze, A.J.; Salomon, A.; Ginley, D.S.; Tillmann, H.H.; Hörhold, H.-H. Polymer–perylene diimide heterojunction solar cells. Appl. Phys. Lett. **2002**, *81* (16), 3085–3087.
30. Sariciftci, N.S.; Braun, D.; Zhang, C.; Srdanov, V.I.; Heeger, A.J.; Stucky, G.; Wudl, F. Semiconducting polymer–buckminsterfullerene heterojunctions—Diodes, photodiodes, and photovoltaic cell. Appl. Phys. Lett. **1993**, *62* (6), 585–587.
31. Yu, G.; Gao, J.; Hummelen, J.C.; Wudl, F.; Heeger, A.J. Polymer photovoltaic cells—Enhanced efficiencies via a network of internal donor–acceptor heterojunctions. Science **1995**, *270* (5243), 1789–1791.
32. Brabec, C.J.; Shaheen, S.E.; Fromherz, T.; Padinger, F.; Hummelen, J.C.; Dhanabalan, A.; Janssen, R.A.J.; Sariciftci, N.S. Organic photovoltaic devices produced from conjugated polymer/methanofullerene bulk heterojunctions. Synth. Met. **2001**, *121* (1–3), 1517–1520.
33. Brabec, C.J.; Sariciftci, N.S.; Hummelen, J.C. Plastic solar cells. Adv. Funct. Mater. **2001**, *11* (1), 15–26.
34. Brabec, C.J.; Zerza, G.; Cerullo, G.; De Silvestri, S.; Luzzati, S.; Hummelen, J.C.; Sariciftci, N.S. Tracing photoinduced electron transfer process in conjugated polymer/fullerene bulk heterojunction in real time. Chem. Phys. Lett. **2001**, *340*, 232–236.
35. Nogueira, A.F.; Montanari, I.; Nelson, J.; Durrant, J.R.; Winder, C.; Sariciftci, N.S.; Brabec, C.J. Charge recombination in conjugated polymer/fullerene blended films studied by transient absorption spectroscopy. J. Phys. Chem., B **2003**, *107*, 1567–1573.
36. Shaheen, S.E.; Brabec, C.J.; Sariciftci, N.S.; Padinger, F.; Fromherz, T.; Hummelen, J.C. 2.5% Efficient organic plastic solar cells. Appl. Phys. Lett. **2001**, *78* (6), 841–843.
37. Geens, W.; Shaheen, S.E.; Wessling, B.; Brabec, C.J.; Poortmans, J.; Serdar Sariciftci, N. Dependence of field-effect hole mobility of PPV-based polymer films on the spin-casting solvent. Org. Electron. **2002**, *3* (3–4), 105–110.
38. Mihailetchi, V.D.; van Duren, J.K.J.; Blom, P.W.M.; Hummelen, J.C.; Janssen, R.A.J.; Kroon, J.M.; Rispens, M.T.; Verhees, W.J.H.; Wienk, M.M. Electron transport in a methanofullerene. Adv. Funct. Mater. **2003**, *13* (1), 43–46.
39. Padinger, F.; Rittberger, R.S.; Sariciftci, N.S. Effects of postproduction treatment on plastic solar cells. Adv. Funct. Mater. **2003**, *13* (1), 85–88.
40. Arias, A.C.; Corcoran, N.; Banach, M.; Friend, R.H.; MacKenzie, J.D.; Huck, W.T.S. Vertically segregated polymer-blend photovoltaic thin-film

structures through surface-mediated solution processing. Appl. Phys. Lett. **2002**, *80* (10), 1695–1697.
41. Huynh, W.U.; Dittmer, J.J.; Alivisatos, A.P. Hybrid nanorod-polymer solar cells. Science (Wash. D.C.) **2002**, *295* (5564), 2425–2427.
42. O'Regan, B.; Grätzel, M. A low cost, high efficiency solar call based on dye sensitized colloidal TiO2 films. Nature **1991**, *353*, 737–740.
43. Krüger, J.; Plass, R.; Grätzel, M.; Matthieu, H.-J. Improvement of the photovoltaic performance of solid-state dye-sensitized device by silver complexation of the sensitizer *cis*-bis(4,4′-dicarboxy-2,2′bipyridine)-bis(isothiocyanato) ruthenium(II). Appl. Phys. Lett. **2002**, *81* (2), 367–369.
44. O'Regan, B.; Lenzmann, F.; Muis, R.; Wienke, J. A solid-state dye-sensitized solar cell fabricated with pressure-treated P_{25}–TiO_2 and CuSCN: Analysis of pore filling and IV characteristics. Chem. Mater. **2002**, *14*, 5023–5029.
45. Lenzmann, F.; O'Regan, B.; Wienke, J.; Huisman, C.; Reijnen, L.; Goossens, A. Surface photovoltage measurements: A useful tool for the detection of electron injection processes in extremely thin absorber (ETA) solar cells. Physica E **2002**, *14*, 233–236.
46. Nozik, A.J. Quantum dot solar cells. Physica E **2002**, *14*, 115–120.
47. Arango, A.C.; Carter, S.A.; Brock, P.J. Charge transfer in photovoltaics consisting of interpenetrating networks of conjugated polymer and TiO_2 nanoparticles. Appl. Phys. Lett. **1999**, *74* (12), 1698–1700.
48. van Hal, P.A.; Wienk, M.M.; Kroon, J.M.; Verhees, W.J.H.; Sloof, L.H.; van Gennip, W.J.H.; Jonkheijm, P.; Janssen, R.A.J. Photoinduced electron transfer and photovoltaic response of a MDMO-PPV:TiO_2 bulk-heterojunction. Adv. Mater. **2003**, *15* (2), 118–121.
49. Lee, S.-J.E.; Chakraborty, A.K. Sequence dependence of polymer dynamics in quenched disordered media: Weak attraction facilitates transport. J. Chem. Phys. **2002**, *117* (23), 10869–10876.
50. Israelachvili, J.N. *Intermolecular and Surface Forces*, 2nd Ed.; Academic Press, Inc.: San Diego, 1997.
51. Coakley, K.M.; Liu, Y.; McGehee, M.D.; Frindell, K.L.; Stucky, G.D. Infiltrating semiconducting polymers into self-assembled mesoporous titania films for photovoltaic applications. Adv. Funct. Mater. **2003**, *13* (4), 301–306.
52. Adams, E.W.; Bruchez, M.P., Jr. Surface Modified Semiconductive and Metallic Nanoparticles Having Enhanced Dispersibility in Aqueous Media. In *PCT Int. Appl. Wo*; Quantum Dot Corporation, USA, 2002.
53. Firestone, M.A.; Csensits, R.; Seifert, S. Polymer-grafted lipid-based complex fluids as scaffolding for the formation of heterostructured nanoparticle arrays. Polym. Prepr. (Am. Chem. Soci., Div. Polym. Chem.) **2002**, *43* (2), 446–447.
54. Torimoto, T. Preparation of monodisperse semiconductor nanoparticles by the size selective photoetching technique and their photoelectrochemical properties. Kagaku Kogyo **2002**, *53* (7), 522–527.
55. Beard, M.C.; Turner, G.M.; Schmuttenmaer, C.A. Size-dependent photoconductivity in CdSe nanoparticles as measured by time-resolved terahertz spectroscopy. Nano Lett. **2002**, *2* (9), 983–987.
56. Yan, C.; Sun, L.; Fu, X.; Liao, C. Preparation of CdS/ZnO core/shell structured nanoparticles by hydrothermal method. Mater. Res. Soc. Symp. Proc. **2002**, *692*, 549–553. (Progress in Semiconductor Materials for Optoelectronic Applications, 2001).
57. Kamiya, I.; Asami, H. Synthesis and functionalization of semiconductor nanoparticles by colloid chemistry. Kagaku Sochi **2002**, *44* (5), 41–45.
58. Baranov, I.; Hakansson, P.; Kirillov, S.; Kopniczky, J.; Novikov, A.; Obnorskii, V.; Pchelintsev, A.; Quist, A.P.; Torzo, G.; Yarmiychuk, S.; Zennaro, L. Desorption of nanoclusters (2–40 nm) from nanodispersed metal and semiconductor layers by swift heavy ions. Nucl. Instrum. Methods Phys. Res., B Beam Interact. Mater. Atoms **2002**, *193*, 798–803.
59. Kowshik, M.; Deshmukh, N.; Vogel, W.; Urban, J.; Kulkarni, S.K.; Paknikar, K.M. Microbial synthesis of semiconductor CdS nanoparticles, their characterization, and their use in the fabrication of an ideal diode. Biotechnol. Bioeng. **2002**, *78* (5), 583–588.
60. Monteiro, O.C.; Trindade, T. Nanochemistry of semiconductors. Química (Lisb.) **2002**, *84*, 49–54.
61. Kawa, M. Semiconductor Nanoparticles. In *Jpn. Kokai Tokkyo Koho, Jp.*; Mitsubishi Chemical Corp., Japan, 2002.
62. Tan, W.; Santra, S.; Zhang, P.; Tapec, R.; Dobson, J. Silica-Coated Nanoparticles. In *PCT Int. Appl. Wo.*; University of Florida, USA, 2001.
63. Crooks, R.M.; Chechik, V.; Lemon, B.I., III; Sun, L.; Yeung, L.K.; Zhao, M. Synthesis, Characterization, and Applications of Dendrimer-Encapsulated Metal and Semiconductor Nanoparticles. In *Metal Nonparticles: Synthesis, Characterization, and Applications;* Feldheim, D.L., Foss, C.A., Jr., Eds.; Dekker: New York, 2002; 261–296.
64. Hong, Y.-K.; Kim, H.; Lee, G.; Kim, W.; Park, J.-I.; Cheon, J.; Koo, J.-Y. Controlled two-dimensional distribution of nanoparticles by spin-coating method. Appl. Phys. Lett. **2002**, *80* (5), 844–846.

65. Mirkin, C.A.; Letsinger, R.L.; Mucic, R.C.; Storhoff, J.J.; Elghanian, R. Nanoparticles Having Oligonucleotides Attached Thereto and Uses in Assays. In *U.S. Us*; Nanosphere, Inc., USA, 2002.
66. Forster, S.; Plantenberg, T. From self-organizing polymers to nanohybrid and biomaterials. Angew. Chem. Int. Ed. **2002**, *41* (5), 688–714.
67. Goldman, E.R.; Mattoussi, H.; Tran, P.T.; Anderson, G.P.; Mauro, J.M. Bioconjugates of luminescent CdSe–ZnS quantum dots with engineered recombinant proteins: Novel self-assembled tools for biosensing. Mater. Res. Soc. Symp. Proc. **2001**, *642*, J2.8/1–J2.8/6. (Semiconductor Quantum Dots II).
68. van der Boom, T.; Hayes, R.T.; Zhao, Y.; Bushard, P.J.; Weiss, E.A.; Wasielewski, M.R. Charge transport in photofunctional nanoparticles self-assembled from zinc 5,10,15,20-tetrakis(perylene-diimide)porphyrin building blocks. J. Am. Chem. Soc. **2002**, *124* (32), 9582–9590.
69. Kovtyukhova, N.I.; Martin, B.R.; Mbindyo, J.K.N.; Mallouk, T.E.; Cabassi, M.; Smith, P.; Razavi, B.; Mayer, T.S. Layer-by-layer assembly of nanotubes and nanofilms from nanoparticles and polymer blocks for electronic applications. NATO Sci. Ser., II: Math. Phys. Chem. **2002**, *57*, 147–162. (Frontiers of Multifunctional Nanosystems).
70. Zhang, H.-M.; Wang, Z.-J.; Zhang, L.-G.; Yuan, J.-S. Synthesis and study of CdS nanoparticles by self-assembled monolayer. Faguang Xuebao **2002**, *23* (4), 369–372.
71. Pacholski, C.; Kornowski, A.; Weller, H. In *Self-Assembly of ZnO: From Nanodots to Nanorods*, Abstracts of Papers, 224th ACS National Meeting, Boston, MA, United States, August 18–22, 2002; 2002, PHYS-374.
72. Cresce, A.V.; Ali, H.A.; Iliadis, A.A.; Lee, U.; Kofinas, P. Synthesis and characterization of self-assembled piezoelectric ZnO nanostructures and integration with Si processing. Mater. Res. Soc. Symp. Proc. **2002**, *728*, 115–120. (Functional Nanostructured Materials through Multiscale Assembly and Novel Patterning Techniques).
73. Sun, L.-D.; Fu, X.-F.; Qian, C.; Liao, C.-S.; Yan, C.-H. Synthesis of core/shell structural CdS/ZnO nanoparticles by hydrothermal method. Gaodeng Xuexiao Huaxue Xuebao **2001**, *22* (6), 879–882.
74. Manna, L.; Scher, E.C.; Alivisatos, A.P. Synthesis of soluble and processable rod-, arrow-, teardrop-, and tetrapod-shaped CdSe nanocrystals. J. Am. Chem. Soc. **2000**, *122* (51), 12700–12706.
75. Gerion, D.; Pinaud, F.; Williams, S.C.; Parak, W.J.; Zanchet, D.; Weiss, S.; Alivisatos, A.P. Synthesis and properties of biocompatible water-soluble silica-coated CdSe/ZnS semiconductor quantum dots. J. Phys. Chem., B **2001**, *105* (37), 8861–8871.
76. Rafeletos, G.; Norager, S.; O'Brien, P. Evidence for the chemical nature of capping in CdSe nanoparticles prepared by thermolysis in tri-*n*-octylphosphine oxide from P-edge EXAFS spectroscopy. J. Mater. Chem. **2001**, *11* (10), 2542–2544.
77. Chen, C.-C.; Chao, C.-Y.; Lang, Z.-H. Simple solution-phase synthesis of soluble CdS and CdSe nanorods. Chem. Mater. **2000**, *12* (6), 1516–1518.
78. Anikin, K.V.; Melnik, N.N.; Simakin, A.V.; Shafeev, G.A.; Voronov, V.V.; Vitukhnovsky, A.G. Formation of ZnSe and CdS quantum dots via laser ablation in liquids. Chem. Phys. Lett. **2002**, *366* (3,4), 357–360.
79. Sun, L.; Xu, B.; Fu, X.; Wang, M.; Qian, C.; Liao, C.; Yan, C. Carboxylic-containing copolymer as template to prepare CdS, ZnS and doped nanoparticles. Sci. China, Ser. B: Chem. **2001**, *44* (1), 23–30.
80. Limmer, S.J.; Hubler, T.L.; Cao, G. Nanorods of various oxides and hierarchically structured mesoporous silica by sol-gel electrophoresis. J. Sol-Gel Sci. Technol. **2003**, *26* (1–3), 577–581.
81. Schneider, J.J. Magnetic core/shell and quantum-confined semiconductor nanoparticles via chimie douce organometallic synthesis. Adv. Mater. (Weinh., Ger.) **2001**, *13* (7), 529–533.
82. Chen, Q.; Zhou, W.; Du, G.; Peng, L.-M. Trititanate nanotubes made via a single alkali treatment. Adv. Mater. **2002**, *14* (17), 1208.
83. Yao, B.D.; Chan, Y.F.; Zhang, W.F.; Yang, Z.Y.; Wang, N. Reaction mechanism of TiO_2 nanotubes. Appl. Phys. Lett. **2003**, *82* (2), 281.
84. Li, D.; Xia, Y. Fabrication of titania nanofibers by electrospinning. Nano Lett. **2003**, *3* (4), 555–560.
85. Xe, Y.; Yan, P.; Lu, J.; Wang, W.; Qian, Y. A safe low temperature route to InAs nanofibers. Chem. Mater. **1999**, *11*, 2619–2622.
86. Gao, S.; Xie, Y.; Lu, J.; Du, G.; He, W.; Cui, D.; Huang, B.; Jiang, M. Mild benzene-thermal route to GaP nanorods and nanospheres. Inorg. Chem. **2002**, *41* (7), 1850–1854.
87. Li, Y.; Liao, H.; Ding, Y.; Fan, Y.; Zhang, Y.; Qian, Y. Solvothermal elemental direct reaction to CdE (E=S, Se, Te). Semiconductor nanorod. Inorg. Chem. **1999**, *38* (7), 1382–1387.
88. Jiang, Y.; Wu, Y.; Mo, X.; Yu, W.; Xie, Y.; Qian, Y. Elemental solvothermal reaction to produce ternary semiconductor cuInE2 (E=S, Se) nanorods. Inorg. Chem. **2000**, *39* (14), 2964–2965.
89. Wang, X.; Gao, P.; Li, J.; Summers, C.J.; Wang, Z.L. Rectangular porous ZnO–ZnS nanocables and ZnS nanotubes. Adv. Mater. **2002**, *14* (23), 1732.

90. Vayssieres, L. Growth of arrayed nanorods and nanowires of ZnO from aqueous solutions. Adv. Mater. **2003**, *15* (5), 464–466.
91. Könenkamp, R.; Boedecker, K.; Lux-Steiner, M.C.; Poschenrieder, M.; Zenia, F.; Levy-Clement, C.; Wagner, S. Thin film semiconductor deposition on free-standing ZnO columns. Appl. Phys. Lett. **2000**, *77* (16), 2575.
92. Lao, J.Y.; Huang, J.Y.; Wang, D.Z.; Ren, Z.F. ZnO nanobridges and nanonails. Nano Lett. **2003**, *3* (2), 235–238.
93. Tang, C.; Fan, S.; de la Chapelle, M.L.; Dang, H.; Li, P. Synthesis of gallium phosphide nanorods. Adv. Mater. (Weinh., Ger.) **2000**, *12* (18), 1346–1348.
94. Dai, Z.R.; Pan, Z.W.; Wang, Z.L. Gallium oxide nanoribbons and nanosheets. J. Phys. Chem., B **2002**, *106*, 902–904.
95. Wang, X.; Li, Y. Rational synthesis of a-MnO_2 single-crystal nanorods. Chem. Commun. (Camb., U. K.) **2002**, (7), 764–765.
96. Liao, H.-W.; Wang, Y.-F.; Liu, X.-M.; Li, Y.-D.; Qian, Y.-T. Hydrothermal preparation and characterization of luminescent CdWO4 nanorods. Chem. Mater. **2000**, *12* (10), 2819–2821.
97. Kim, F.; Kwan, S.; Akana, J.; Yang, P. Langmuir–Blodgett nanorod assembly. J. Am. Chem. Soc. **2001**, *123* (18), 4360–4361.
98. Kwan, S.; Kim, F.; Akana, J.; Yang, P. Synthesis and assembly of $BaWO_4$ nanorods. Chem. Commun. (Camb., U. K.) **2001**, (5), 447–448.
99. Chen, X.; Sun, X.-M.; Li, Y. Self-assembling vanadium oxide nanotubes by organic molecular templates. Inorg. Chem. **2002**, *41* (17), 4524.
100. Tenne, R. Inorganic fullerene-like materials and inorganic nanotubes from layered metal dichalcogenides and other compounds. Abstr. Pap.-Am. Chem. Soc. **2001**, *221st*. INOR-643.
101. Tenne, R. Fullerene-like materials and nanotubes from inorganic compounds with a layered (2-D) structure. Colloids Surf., A: Physicochem. Eng. Asp. **2002**, *208* (1–3), 83–92.
102. Zak, A.; Feldman, Y.; Lyakhovitskaya, V.; Leitus, G.; Popovitz-Biro, R.; Wachtel, E.; Cohen, H.; Reich, S.; Tenne, R. Alkali metal intercalated fullerene-like MS_2 (M=W, Mo) nanoparticles and their properties. J. Am. Chem. Soc. **2002**, *124* (17), 4747–4758.
103. Zhu, Y.Q.; Hsu, W.K.; Firth, S.; Terrones, M.; Clark, R.J.H.; Kroto, H.W.; Walton, D.R.M. Nb-doped WS_2 nanotubes. Chem. Phys. Lett. **2001**, *342* (1,2), 15–21.
104. Hickey, S.G.; Riley, D.J.; Tull, E.J. Photoelectrochemical studies of CdS nanoparticle modified electrodes: Absorption and photocurrent investigations. J. Phys. Chem., B **2000**, *104* (32), 7623–7626.
105. Hupp, J.T.; Gaal, D.A.; Leytner, S. In *Kinetics, Energetics, and Dynamics of Inorganic-Dye/Semiconductor-Nanoparticle Electron Transfer Reactions*, Abstracts of Papers, 222nd ACS National Meeting, Chicago, IL, United States, August 26–30, 2001; **2001**, CHED-286.
106. Huynh, W.U.; Dittmer, J.J.; Libby, W.C.; Whiting, G.L.; Alivisatos, A.P. Controlling the morphology of nanocrystal-polymer composites for solar cells. Adv. Funct. Mater. **2003**, *13* (1), 73–79.
107. Kamat, P.V. Photophysical, photochemical and photocatalytic aspects of metal nanoparticles. J. Phys. Chem., B **2002**, *106* (32), 7729–7744.
108. Torimoto, T.; Ohtani, B. Special articles: Nanostructured materials for electrochemistry. Preparation and immobilization of size-quantized semiconductor nanoparticles and their photoelectrochemical properties. Electrochemistry (Tokyo, Japan) **2001**, *69* (11), 866–871.

Polyelectrolyte–Surfactant Complex Nanoparticles

Hans-Peter Hentze
University of Delaware, Newark, Delaware, U.S.A.

INTRODUCTION

A variety of different colloidal systems can be produced by interaction between polyelectrolytes and oppositely charged surfactants. These complexes have attracted considerable interest because of their importance for both science and industrial applications.

By interaction of two oppositely charged polyelectrolytes, polyelectrolyte–polyelectrolyte complexes (PECs, or so-called symplexes) are formed. These PECs can be loose aggregates or well-defined nanoparticles.[1] In a similar way polyelectrolytes form complexes with oppositely charged surfactants and polyelectrolyte–surfactant complexes (PSCs) are obtained.[2,3] Here the mechanism of complex formation is not only based on electrostatic interactions and entropic factors, but also on the hydrophobic interaction between the surfactant tails. At surfactant concentrations below the critical micelle concentration (CMC), complex formation of polyelectrolytes with oppositely charged surfactants occurs. With increasing surfactant chain lengths, the formation of surfactant patches along the polyelectrolyte chain can be expected, because owing to hydrophobic interactions the preferred complexation binding site is the next closest position to an occupied site. This results in a zipperlike mechanism of polyelectrolyte–surfactant complex formation.[4] Typically three-dimensional networks form at concentrations above the CMC, where polyelectrolyte chains bridge discrete micelles without penetrating them.[5,6] At even higher surfactant concentrations, when all charges of the polyion are neutralized, polyelectrolyte–surfactant complexes usually precipitate from solution.[2,3,7] These bulk materials have a liquid crystalline-like mesostructure and display long-range order on the nanometer scale. The interaction between polyions and oppositely charged surfactants is a cooperative process in which the ionic head groups of the surfactant bind to the polyion, while the hydrophobic surfactant tails segregate into a hydrophobic microphase. Similar to amphiphilic block copolymers, this microphase separation occurs without macroscopic demixing.

One problem with most such polyelectrolyte–surfactant complexes is that they are usually water-insoluble solids or semisolids and not suitable for particle applications. Within the last few years, different techniques of PSC nanoparticle synthesis have been developed to overcome this problem. These complexes are, unlike PSC bulk materials, redispersable in water or organic solvents as discrete nanoparticles without dissociation. For synthesis of PSC nanoparticles, polyelectrolytes that exhibit additional molecular entities have to be used, which provide solubility even after stoichiometric complexation and neutralization of all charges. In this way, particles are obtained that exhibit spherical core–shell structures, cylindrical or disk-shaped morphologies, or hollow spheres like vesicles. Whereas some of these morphologies appear to be equilibrium structures, others are nonequilibrium structures. The size, solubility, and morphology of PSC nanoparticles depend strongly on the molecular structure of the polyelectrolyte, the block ionomer lengths, the block length ratios, the structure and type of surfactant, and sometimes also on the way of preparation. Potential applications for this relatively new class of materials are drug and gene delivery, amphiphilic particles, and depot systems.[8,9]

COMPLEXATION OF DOUBLE-HYDROPHILIC BLOCK COPOLYMERS—VARIOUS NANOPARTICLE MORPHOLOGIES

A very interesting family of polymers are double-hydrophilic block copolymers.[10] These consist of two covalently linked hydrophilic polymer block segments, which are chemically not equivalent, e.g., poly(ethylene oxide)-b-poly(sodium methacrylate). Double-hydrophilic block copolymers that contain a polymer block that is a polyelectrolyte are called block ionomers.

For the first time, redispersable nanoparticles of block ionomer–surfactant complexes have been obtained by complexation of poly(ethylene oxide)-b-poly(sodium methacrylate) (PEO-b-PMANa) with cetylpyridinium cations.[11] The segment lengths of the block ionomer were 176 and 186 repeating units for PEO and PMANa, respectively. Different aggregate morphologies could be distinguished within three different surfactant concentration regions: At molar surfactant concentrations below the concentration of anionic polyion charges, the surfactant

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120014189

cetylpyridinium bromide (C_{16}PyBr) binds electrostatically to the polyion and soluble complexes with negative ζ-potentials were observed. Also, the stoichiometric and electrostatic neutral ($\zeta = 0$) complexes were soluble and formed monodisperse nanoparticles with a diameter of 67 nm, as shown by dynamic light scattering (DLS) and transmission electron microscopy (TEM). A TEM micrograph is shown in Fig. 1. These discrete PSC complexes were first believed to exhibit a core–shell structure, where the core is formed by C_{16}PyBr neutralized polyions, and the shell of PEO block segments. Later studies suggested a vesicular morphology.[12] At higher surfactant concentrations excess surfactant is incorporated into the particles, resulting in a positive ζ-potential. The observed particles combine the properties of amphiphilic block copolymers and polyelectrolyte surfactant complexes, as they spontaneously form stable, water-soluble, and microphase separated aggregates.

Similar to mixtures of cationic and anionic single-tail surfactants,[13] spontaneous formation of vesicles is observed for the complexation of PEO-b-PMANa block ionomers with alkyltrimethylammonium surfactants [e.g., tetradecylammonium bromide (TTAB) or cetyltrimethylammonium bromide (CTAB)].[12] In contrast, complexation of the homopolymer PMANa with the same surfactants resulted in the formation of water-insoluble bulk materials. The wall of the PSC vesicles is composed of surfactant-neutralized polyion surrounded by the hydrophilic PEO shell, which sterically stabilizes the complexes in aqueous dispersions. The vesicular solutions are transparent or slightly opalescent. Vesicle diameters varied between 85 and 120 nm and increased with increasing chain length of the surfactant. The obtained dispersions are very stable and efficiently encapsulate and retain hydrophilic molecules in the internal aqueous volume.

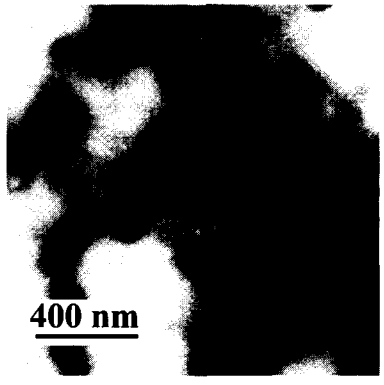


Fig. 1 TEM micrograph of vesicles formed by the complex of PEO-b-PMANa with C_{16}PyBr. (From Ref. [11].)

The influence of the block ionomer block lengths and the surfactant structure on the solubility of PSC nanoparticles was also systematically investigated.[14] Therefore, complexation of a variety of PEO-b-PMANa block copolymers was performed with single-tail, double-tail, and triple-tail tetraalkylammonium surfactants. It was found that a minimum block length of PEO and a high enough block length ratio of PEO/PMANa are necessary to induce sufficient sterical stabilization in aqueous dispersions. Complex formation with stoichiometric amounts of the single-tail surfactants CTAB or C_{16}PyBr resulted in rather small and monodisperse aggregates. These systems are believed to be in dynamic equilibrium with surfactant monomers in the dispersion. The critical aggregation concentration (CAC) was two to three orders of magnitude lower than the corresponding CMC. In contrast, the aggregate size of complexes formed with double-tailed surfactants [e.g., didodecyldimethylammonium bromide (DDAB)] depended strongly on the way of preparation. This nonequilibrium behavior can be explained by the formation of preformed, membranelike surfactant aggregates, which is typical for double-tail surfactants. As is the case for single-tail surfactants, for triple-tail surfactants the formation of the aggregates is primarily controlled by the block ionomer structure and composition of the mixture. Different core morphologies can be expected for complexation with single- and triple-tail surfactants owing to the rather different packing parameters.

Not only linear block copolymers, but also grafted copolymers were used for PSC nanoparticle formation.[15] Unlike the vesicular morphologies formed by linear block copolymers, core–shell particles were obtained by complexation of poly(ethylene oxide)-g-poly(ethylene imine) (PEO-g-PEI) with anionic surfactants [e.g., sodium dodecyl sulfate (SDS) or sodium tetradecyl sulfate (TDS)]. In these micelle-like aggregates, apparently the hydrophobic core from the surfactant-neutralized PEI block is surrounded by the hydrophilic corona of PEO chains stabilizing the particles in aqueous media. The observed particle sizes, as determined by DLS and confirmed by TEM, are about 40–50 nm. Therefore, their size is comparable with micelle-like aggregates of amphiphilic block copolymers. The CAC was determined by fluorescence measurements using pyrene as a probe. CAC values were about two magnitudes lower than the corresponding CMC values of the pure surfactant. As the size of the obtained PSC particles is independent of the way of preparation, the complexes appear to be in equilibrium with unbound surfactants. With increasing chain length of the anionic surfactants, excess surfactant was incorporated in the complexes, resulting in particles with negative ζ-potentials.

Spherical core–shell morphologies can also be obtained by complexation of poly(ethylene oxide)-b-poly(ethylene

imine) containing branched poly(ethylene imine) segments (PEO-b-PEI_{br}) with the surfactant dodecanoic acid (C_{12}).[16] In contrast, prolate particles can be obtained by complexation of the linear block copolymer PEO-b-PEI_{li}. Even more elongated are the needlelike particles formed by C_{12} and a PEO-b-PEI block copolymer with a cyclic architecture of the PEI segment (PEO-b-PEI_{cy}). All these particles formed by C_{12} and PEO-b-PEI block copolymers with a linear, branched, or cyclic PEI architecture exhibit a core–shell structure, as shown by atomic force microscopy (AFM). The core consists of poly(ethylene imine) dodecanoate surrounded by a corona of poly(ethylene oxide). All particles investigated in this study have an average size of about 200 nm (Fig. 2). Small-angle X-ray scattering (SAXS) of dried dispersions revealed that the complex formed with PEO-b-PEI_{cy} has a lamellar morphology, while the ones of PEO-b-PEI_{li} and PEO-b-PEI_{br} show a lamellar-*within*-lamellar morphology.

pH-sensitive core–shell particles have been obtained in aqueous dispersions of poly(ethylene oxide)-b-poly(L-lysine) (PEO-b-PLL) complexes with retinoic acid.[17] Because of its amphiphilic properties, this vitamin A analogue can be considered as a surfactant. The complexes were characterized by SAXS as bulk material, as well as in aqueous dispersions. The diffractograms indicate that the cores of these complexes exhibit a smectic-A-like structure. By circular dichroism, an α-helical conformation of L-lysine chains was found for pH values above 9. It was assumed that this conformation is stabilized by the PEO shell, as well as by protecting retinoate molecules. Below a pH of 3.7, a random coil structure is present, while mixtures of random coils and α-helices form in the intermediate pH range. The complexes can be considered as a model for pH-sensitive drug-carrier systems.

Nanoparticle dispersions of polyelectrolyte–surfactant complexes have also been made from polyelectrolyte homopolymers.[18] Nonstoichiometric complexes containing excess polyelectrolyte or surfactant are generally soluble in water, and sometimes they form well-defined, discrete aggregates. Especially, PSCs of biopolymers, such as polysaccharides or polypeptides, have been investigated for drug and gene delivery purposes. As their stoichiometric, electroneutral complexes are not redispersable and water insoluble they are not discussed in detail in this entry.

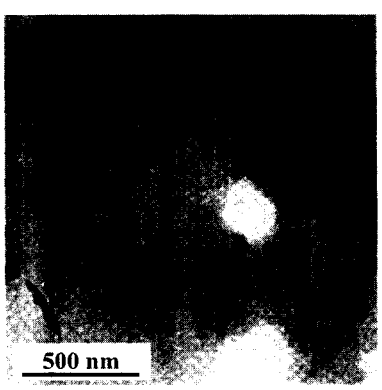


Fig. 2 Dried dispersions of the complex of PEO–PEI_{br} with C12 form needlelike particles. (From Ref. [16].)

"REVERSE" CORE–SHELL STRUCTURES BY COMPLEXATION OF HYDROPHILIC–HYDROPHOBIC BLOCK IONOMERS

In contrast to the core–shell structure of complexes from double-hydrophilic block ionomers, complexation of hydrophilic–hydrophobic block copolymers results in "reverse" morphologies—here the nonionic, hydrophobic block segments form the core, which is surrounded by a surfactant neutralized polyion shell. For this reason the particles can be redispersed only in solvents that are also solvents of the polyelectrolyte homopolymer complex. A major difference compared to double-hydrophilic block copolymers is the aggregation of the polyion before addition of surfactant and complexation. Micelle-like aggregates are formed in aqueous solution owing to microphase separation of the hydrophobic polymer segments. This factor might limit the possibilities of this approach to form equilibrium structures.

One example is the synthesis of nanoparticles from polystyrene-b-poly(sodium acrylate) (PS-b-PAA^-) and *N*-cetylpyridinium cations ($C_{16}Py^+$), which precipitate from aqueous solution at stoichiometric mixing ratios of surfactant and polyion units.[19] The formed precipitate is insoluble in tetrahydrofuran and toluene, which are good solvents for PS but nonsolvents for $PAA^-/C_{16}Py^+$ complexes. Instead, chloroform, a good solvent for $PAA^-/C_{16}Py^+$, could be used for redispersion. This clearly indicates a core–shell structure with a polystyrene core surrounded by surfactant neutralized polyion segments that stabilize the particles in solution.

Similar results were obtained for complexes from poly(α-methylstyrene)-b-poly(*N*-ethly-4-vinylpyridinium bromide) (PMeS-b-PE4VP) and sodium bis(2-ethylhexyl)sulfosuccinate (AOT).[20] The precipitated complexes can incorporate small excess of surfactant. Preliminary studies showed that PSC formed by the homopolymer PE4VP and AOT were soluble in various organic solvents without dissociation of the complex, e.g., cyclohexane, chloroform, and acetone. Neither of these solvents dissolves the block ionomer PMeS-b-PE4VP, but all of them are good solvents for the complexes of the block ionomer with AOT. Again, this clearly hints on the

described "reverse" core–shell morphology. The particle size of the redispersed PSCs in hexane/cosolvent mixtures is about 70 nm in diameter. The spherical morphology and the size of the particles were confirmed by TEM.

COMPLEXATION OF POLYAMPHOLYTES

Polyampholytes are polyelectrolytes that contain both anionic and cationic repeating units. Complexes of ternary polyampholytes with cationic, anionic, and neutral repeating units with perfluorododecanoic acid (FC_{12}) have been synthesized.[21] The obtained nanoparticles are highly fluorinated and stable in aqueous dispersions. By dynamic light scattering, it was shown that discrete particles with hydrodynamic diameters of 20–35 nm were formed. The diameter and the stability of the dispersions depend strongly on the composition of the polyampholyte. More insight into the particles' morphology was gained by small-angle X-ray scattering. The geometry of the dispersed particles was characterized by indirect Fourier transform method.[22] Disk-shaped aggregates with a diameter of 30 nm were formed by complexes from the polyampholyte with the highest content of cationic repeating units. Complexes of polyampholytes with a lower content of cationic charges formed cylindrical aggregates with a length of about 25 nm. A partially interdigitated arrangement of the surfactant tails was assumed, owing to the small thickness of the disklike aggregates (2.2 nm) and the small diameter of the formed cylinders (3 nm). An overview of the different morphologies of PSC nanoparticles formed by the different polyion precursors is given in Table 1.

HOLLOW SPHERES FROM POLYELECTROLYTE–SURFACTANT COMPLEX NANOPARTICLES

Besides vesicle formation, hollow spheres of PSCs can be also prepared by stepwise complexation and selective dissociation of poly(ethylene imine).[23] First, nonstoichiometric complexes of PEI homopolymer with dodecanoic acid were formed. The particles display a positively charged shell of free ethylene imine units and a PEI–dodecanoic acid core. Second, poly(ethylene oxide)-b-poly(methacrylic acid) block copolymer with a phosphonized carboxylic acid end group was added, resulting in particles with a core of PEI–dodecanoic acid and a shell of a phosphonized block ionomer/PEI complex. By decreasing the pH value below the pK_a value of dodecanoic acid, the PEI–dodecanoic acid core was selectively dissolved.

Table 1 Overview of the different types of polyion precursors used for complexation of oppositely charged surfactants, and the resulting nanoparticle morphologies

Structure of the polyion		Surfactant (reference)	Nanoparticle morphology
Double-hydrophilic block copolymers	PEO-b-PMANa	$C_{16}PyBr$[11]	Vesicles
		DOTAB[12]	
		TTAB[12]	
		CTAB[12]	
	PEO-g-PEI	SDS[15]	Core–shell
		TDS[15]	
		AOT[15]	
	PEO-b-PLL	Retinoate[17]	
	PEO-b-PEI_{br}	C_{12}[16]	
	PEO-b-PEI_{li}	C_{12}[16]	Prolate particles
	PEO-b-PEI_{cy}	C_{12}[16]	Needlelike particles
Hydrophilic–hydrophobic block ionomers	PS-b-PAA^-	$C_{16}PyBr$[19]	"Reverse" core–shell structures
	PMeS-b-PE4VP	AOT[20]	
Polyampholytes	Random copolymers from cationic, anionic and neutral monomers	FC_{12}[21]	Disk-shaped particles and cylinders
Core–shell latex particles	PEO-b-PSS particles	DOTAB[24]	Core–shell
		TTAB[24]	
		CTAB[24]	
		DDAB[24]	

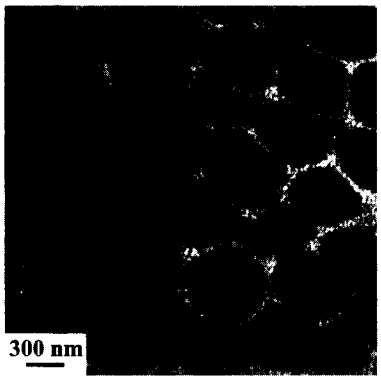


Fig. 3 Micrograph of a microtomed sample of preformed latex particles with a PSS-TTAB core and a PEO shell. (Ref. [24].)

The phosphonate–PEI shell was left unchanged, forming hollow spheres with a diameter of about 400 nm. TEM micrographs reveal a porous structure of the formed PSC shells, which might have formed owing to the high osmotic pressure inside the hollow spheres induced by dissociation of the core.

COMPLEXATION OF PREFORMED CORE–SHELL LATEX PARTICLES

Core–shell particles of PSC complexes have also been made from preformed latex particles.[24] For this purpose core–shell latex particles with a polystyrene core and a poly(ethylene oxide) shell were synthesized by cross-linking emulsion polymerization of styrene, using a PEO macroinitiator system. Subsequent sulfonation of the PS core resulted in particles with a polyelectrolyte core of poly(styrene sulfonate) surrounded by a PEO corona (PEO-b-PSS). Complexation with tetraalkylammonium surfactants [e.g., dodecyltrimethylammonium bromide (DOTAB)] induced microphase separation within the core. Hereby polymer latex particles with a mesostructured core and a diameter of about 400 nm were obtained (Fig. 3). By small-angle X-ray scattering, it was shown that the characteristic length of the formed core mesophase varied between 2 and 4 nm and increased with increasing chain length of the cationic surfactant.

CONCLUSION

Complexation of block ionomers, polyampholytes, or preformed core–shell lattices enables the synthesis of well-defined, monodisperse PSC nanoparticles. Depending on the particular system, these nanoparticles form dispersions in water or organic solvents. In this way, particle applications of stoichiometric PSC have also been made possible.

Compared to bulk polyelectrolyte–surfactant complexes, the synthesis of PSC nanoparticles largely extends the opportunities of structuring nanoparticles. Morphological investigations indicate that self-assembly in these systems mainly depends on the block copolymer architecture. By using different types of polyion precursors, complexation with surfactants results in cylindrical, disk-shaped, prolate, spherical, and vesicular polymer structures. The obtained nanoparticles exhibit a long-range ordered inner mesostructure owing to microphase separation of the surfactant tails from the surfactant-neutralized polyion segments. Potential applications are especially drug and gene release systems.

REFERENCES

1. Buchhammer, H.-M.; Petzold, G.; Lunkwitz, K. Nanoparticles based on polyelectrolyte complexes: Effect of structure and net charge on the sorption capability for solved organic molecules. Colloid Polym. Sci. **2000**, *278*, 841–847.
2. Zhou, S.Q.; Chu, B. Assembled materials: Polyelectrolyte–surfactant complexes. Adv. Mater. **2000**, *12*, 545–556.
3. Thünemann, A.F. Polyelectrolyte–surfactant complexes (synthesis, structure and materials aspects). Prog. Polym. Sci. **2002**, *27*, 1473–1572.
4. Philipp, B.; Dawydoff, W.; Linow, K.-J. Polyelectrolyte complexes. Z. Chem. **1982**, *22*, 1.
5. Jonsson, M.; Linse, P. Polyelectrolyte–macroion complexation. I. Effect of linear charge density, chain length, and macroion charge. J. Chem. Phys. **2001**, *115*, 3406–3418.
6. Macdonald, P.M. ^{2}H NMR and polyelectrolyte–surfactant interactions: From micelles to monolayers to membranes. Colloids Surf., A Physicochem. Eng. Asp. **1999**, *147*, 115–131.
7. Antonietti, M.; Conrad, J.; Thünemann, A.F. Polyelectrolyte–surfactant complexes—A new type of solid, mesomorphous material. Macromolecules **1994**, *27*, 6007–6011.
8. Kabanov, A.V.; Vinogradov, S.V.; Suzdaltseva, Y.G.; Alakhov, V.Y. Water-soluble block polycations as carriers for oligonucleotide delivery. Bioconjug. Chem. **1995**, *6*, 639–643.
9. Kataoka, K.; Togawa, H.; Harada, A.; Yasugi, K.; Matsumoto, T.; Katayose, S. Spontaneous formation of polyion complex micelles with narrow distribution from antisense oligonucleotide and cationic block copolymer in physiological saline. Macromolecules **1996**, *29*, 8556–8557.

10. Cölfen, H. Double-hydrophilic block copolymers: Synthesis and application as novel surfactants and crystal growth modifiers. Macromol. Rapid Commun. **2001**, *22*, 219–252.
11. Bronich, T.K.; Kabanov, A.V.; Kabanov, V.A.; Yu, K.; Eisenberg, A. Soluble complexes from poly (ethylene oxide)-block-polymethacrylate anions and *N*-alkylpyridinium cations. Macromolecules **1997**, *30*, 3519–3525.
12. Kabanov, A.V.; Bronich, T.K.; Kabanov, V.A.; Yu, K.; Eisenberg, A. Spontaneous formation of vesicles from complexes of block ionomers and surfactants. J. Am. Chem. Soc. **1998**, *120*, 9941–9942.
13. Kaler, E.W.; Murthy, A.K.; Rodriguez, B.E.; Zasadzinski, J.A. Spontaneous vesicle formation in aqueous mixtures of single-tailed surfactants. Science **1989**, *245*, 1371–1374.
14. Bronich, T.K.; Popov, A.M.; Eisenberg, A.; Kabanov, V.A.; Kabanov, A.V. Effects of block length and structure of surfactant on self-assembly and solution behavior of block ionomer complexes. Langmuir **2000**, *16*, 481–489.
15. Bronich, T.K.; Cherry, T.; Vinogradov, S.V.; Eisenberg, A.; Kabanov, V.A.; Kabanov, A.V. Self-assembly in mixtures of poly(ethylene oxide)-graft-poly(ethyleneimine) and alkyl sulfates. Langmuir **1998**, *14*, 6101–6106.
16. Thünemann, A.F.; General, S. Poly(ethylene oxide)-b-poly(ethylene imine) dodecanoate complexes: Lamellar-within-lamellar morphologies and nanoparticles. Macromolecules **2001**, *34*, 6978–6984.
17. Thünemann, A.F.; Beyermann, J.; Kukula, H. Poly(ethylene oxide)-b-poly(L-lysine) complexes with retinoic acid. Macromolecules **2000**, *33*, 5906–5911.
18. General, S.; Thünemann, A.F. pH-sensitive nanoparticles of poly(amino acid) dodecanoate complexes. Int. J. Pharm. **2001**, *230*, 11–24.
19. Lysenko, E.A.; Bronich, T.K.; Eisenberg, A.; Kabanov, V.A.; Kabanov, A.V. Block ionomer complexes from polystyrene-block-polyacrylate anions and *N*-cetylpyridinium cations. Macromolecules **1998**, *31*, 4511–4515.
20. Lysenko, E.A.; Bronich, T.K.; Eisenberg, A.; Kabanov, V.A.; Kabanov, A.V. Solution behavior and self-assembly of complexes from poly(a-methylstyrene)-block-poly(*N*-ethyl-4-vinylpyridinium) cations and aerosol OT anions. Macromolecules **1998**, *31*, 4516–4519.
21. Thünemann, A.F.; Wendler, U.; Jaeger, W.; Schnablegger, H. Nanoparticles of polyelectrolyte–surfactant complexes with perfluorododecanoic acid. Langmuir **2002**, *18*, 4500–4504.
22. Glatter, O. A new method for the evaluation of small-angle scattering data. J. Appl. Crystallogr. **1977**, *10*, 415–421.
23. General, S.; Rudloff, J.; Thünemann, A.F. Hollow nanoparticles via stepwise complexation and decomplexation of poly(ethylene imine). Chem. Commun. **2002**, *5*, 534–535.
24. Hentze, H.-P.; Khrenov, V.; Tauer, K. A new approach towards redispersable polyelectrolyte–surfactant complex nanoparticles. Colloid Polym. Sci. **2002**, *280*, 1021–1026.

Polymer Colloids and Their Metallation

Lyudmila M. Bronstein
Indiana University, Bloomington, Indiana, U.S.A.

INTRODUCTION

Polymer colloids with metal nanoparticles is a fast-developing field of nanoscience and nanotechnology as encapsulation of metal inside the polymer—or, vice versa, formation of polymer inside the nanocomposite colloids—show a promise for exciting applications: catalytic, optical, magnetic. Metal or semiconductor nanoparticles can be 1) formed inside the polymer colloids (block copolymer micelles, dendrimers, or other functionalized polymer colloids); 2) positioned on the outer surface of the polymer colloids; or 3) polymer layer can be formed or adsorbed on the preformed nanoparticle surface. This entry discusses the above routes to prepare polymer nanocomposite colloids considering possible advantages and disadvantages of different methods. The major methods to characterize composition and structure of these nanomaterials include those used for polymer colloid characterization and metal nanoparticle assessment. Some methods allow gathering information on both polymer and nanoparticle structure; some methods are very specific for nanoparticle characterization. Transmission electron microscopy (TEM) is widely used to characterize both polymer colloids and metal (semiconductor) nanoparticles; the latter especially provide high electron contrast. High-resolution TEM (HRTEM) is used to determine the sizes of small particles and to characterize their inner structure. X-ray diffraction (XRD) is used for crystalline nanoparticles to determine their structure and mean particle size. Scattering techniques, small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS), can be used both for polymer colloid and nanoparticle assessment. The latter is possible if polymer structure does not change after particle formation so scattering from polymer particles can be subtracted. If this is not the case, anomalous SAXS is employed because it allows one to establish nanoparticle size and particle size distribution independently of changes in the nanostructured polymer. Spectroscopic techniques are often used to characterize nanoparticles because properties are dependent on nanoparticle size (Au, Ag, semiconductor nanoparticles). In the sections below, examples of polymer colloid and nanoparticle characterization are presented.

NANOPARTICLE FORMATION INSIDE POLYMER COLLOIDS

Nanoparticle Formation in Block Copolymer Micelles

Amphiphilic block copolymers form micelles in selective solvents (a good solvent only for one block), yet the size and shape of micelles depend on the block chemical structure, molecular weight of each block, and solvent type.[1,2] Block copolymer micelles can be treated as polymer colloids being in dynamic equilibrium with unimers (individual macromolecules in solution) and with each other, which results in exchange between micelles. Although this exchange is a slow process for block copolymer systems compared to surfactant micelles, this can ensure very useful properties (ability to assemble and disassemble in certain conditions can play very important role for some "delivery" applications) or demonstrate a disadvantage, as it can facilitate metal-species exchange between polymer colloids.

Depending on the structure of the block copolymer, nanoparticles can be formed both in the micelle core (when the core is functionalized, while corona is not) or in the corona, when the core is not functionalized.

Nanoparticle formation in the micelle core

If the block containing the functional groups (able to react with metal compounds, giving complexes or salts) forms the micelle core, it can be loaded with a corresponding metal compound (by incorporating the metal compound in the block copolymer solution) and can further serve as a nanoreactor for nanoparticle formation. In so doing, because the core-forming block is not soluble in a selective solvent, the micelle core can be treated like a quasi solid, thus additionally stabilizing the nanoparticles. Examples of "functional" blocks are polyvinylpyridines [P2(4)VP], polymethacrylic and polyacrylic acids (PMAA and PAA), polybutadiene (PB), polyisoprene (PI), and others. A block containing no functional groups but providing solubility and micelle stability in the solution should form the micelle corona. These can be polystyrene (PS), poly(ethylene oxide) (PEO), polyisobutylene (PIB), etc.

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013732

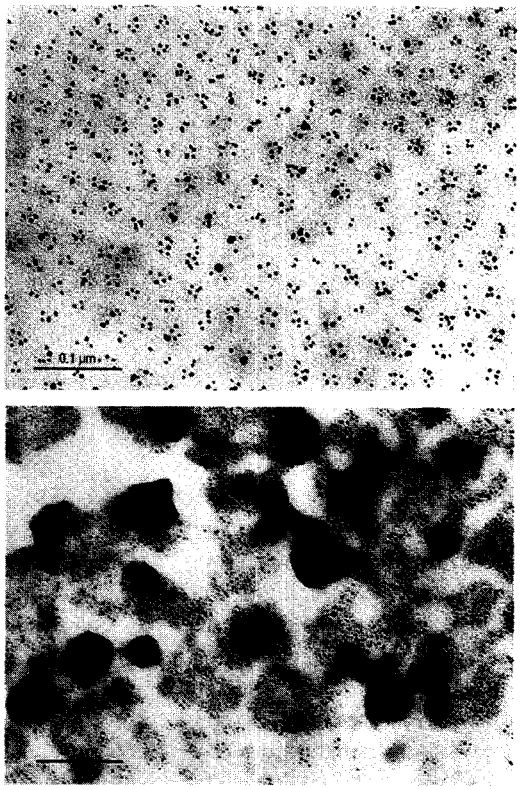


Fig. 1 Electron micrographs of Pd colloids synthesized in PS-*b*-P4VP block copolymer micelles via reduction with hydrazine (top) and $NaBH_4$ (bottom). (From Ref. [9] Copyright 1997 by the American Chemical Society.)

The synthesis of metal or semiconductor nanoparticles in the cores of amphiphilic block copolymer micelles was almost simultaneously reported by several research groups.[3–7] The metal nanoparticle formation in PS-*b*-P4VP block copolymer micelles demonstrated the strong dependence of nanoparticle morphology on the type of reducing agent. When a sluggish reducing agent is used, one nanoparticle per micelle ("cherry-like" morphology) can be formed if there is no exchange between micelles (e.g., the micelle is cross-linked).[8] Fast reduction leads to the formation of many small particles per micelle ("raspberry-like" morphology), which is considered to be preferable for catalytic applications (Fig. 1).[9] Using block copolymer micelle cores as nanoreactors allows synthesis of monometallic and bimetallic nanoparticles, yet bimetallic particle morphology depends on a metal pair,[10] i.e., on the ability of metal species to be reduced in particular conditions. For the Pd–Au pair, core–shell particles are formed with a gold core and a palladium shell. For Pd–Pt pair, cluster-in-cluster particles are obtained. These different morphologies significantly change the catalytic properties of such systems although the nanoparticle sizes are similar.[10]

Co nanoparticles of different sizes and shapes can be prepared either by incorporation of $CoCl_2$ in the PS-*b*-P2VP micelles followed by reduction or by thermal decomposition of $Co_2(CO)_8$ species embedded in the micelle cores.[11] Stable suspensions of superparamagnetic cobalt nanoparticles were also prepared in poly(dimethylsiloxane) (PDMS) carrier fluids in the presence of poly[dimethylsiloxane-*block*-(3-cyanopropyl)methylsiloxane-*block*-dimethylsiloxane] (PDMS-*b*-PCPMS-*b*-PDMS) triblock copolymers as steric stabilizers.[12] Similar to PS-*b*-P2VP, these copolymers formed micelles in toluene and served as nanoreactors for thermal decomposition of the $Co_2(CO)_8$ precursor. The nitrile groups on the PCPMS central blocks are thought to adsorb onto the particle surface, while the PDMS end blocks protrude into the reaction medium to provide steric stability. Adjusting the cobalt-to-copolymer ratio can control the particle size. Transmission electron microscopy shows nonaggregated cobalt nanoparticles with a narrow size distribution and the particles are evenly surrounded with copolymer covering.

The formation of iron oxide particles in cross-linked block copolymer micelles is described in Ref. [13]. A polyisoprene-*block*-poly(2-cinnamoylethyl methacrylate)-*block*-poly(*tert*-butyl acrylate), PI-*b*-PCEMA-*b*-P*t*BA, forms spherical micelles in THF/hexane mixture with 65% volume fraction of the latter. The micelles consist of a PI corona, a solvent-insoluble PCEMA shell, and a P*t*BA core. Their structure is locked in by photo-cross-linking the PCEMA shell to yield nanospheres (Fig. 2). Similar to

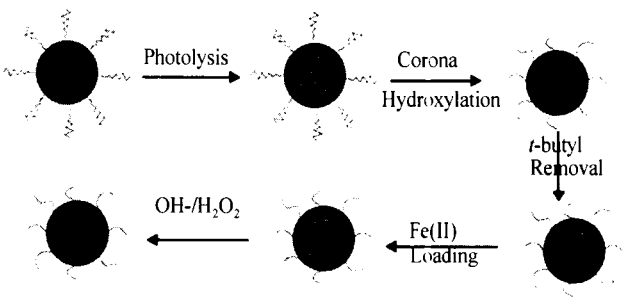


Fig. 2 Preparation of water-dispersible magnetic nanoparticles. Photolysis cross-links the PCEMA shell (gray to dark). The PI corona chains are made water-soluble by hydroxylating the double bonds (wavy lines to free-hand lines). The core is made inorganic compatible by removing the *tert*-butyl groups (light gray to gridded pattern). Soaking the nanospheres in aqueous $FeCl_2$ enables proton exchange by Fe^{2+} (slant to vertical grids) and the Fe^{2+} ions are precipitated and oxidized to yield cubic γ-Fe_2O_3 magnetic particles using NaOH and H_2O_2 (last step). (From Ref. [13] Copyright 2000 by the American Chemical Society.)

core cross-linking, this approach prevents exchange between micelles. The nanospheres were made water-dispersible by hydroxylating the PI double bonds. The core was made compatible with inorganic species by removing the *tert*-butyl groups of P*t*BA. The possibility of using such nanospheres as nanoreactors for inorganic nanoparticle preparation was demonstrated by incorporating iron salt and formation of iron oxide magnetic particles in the cores (Fig. 3).

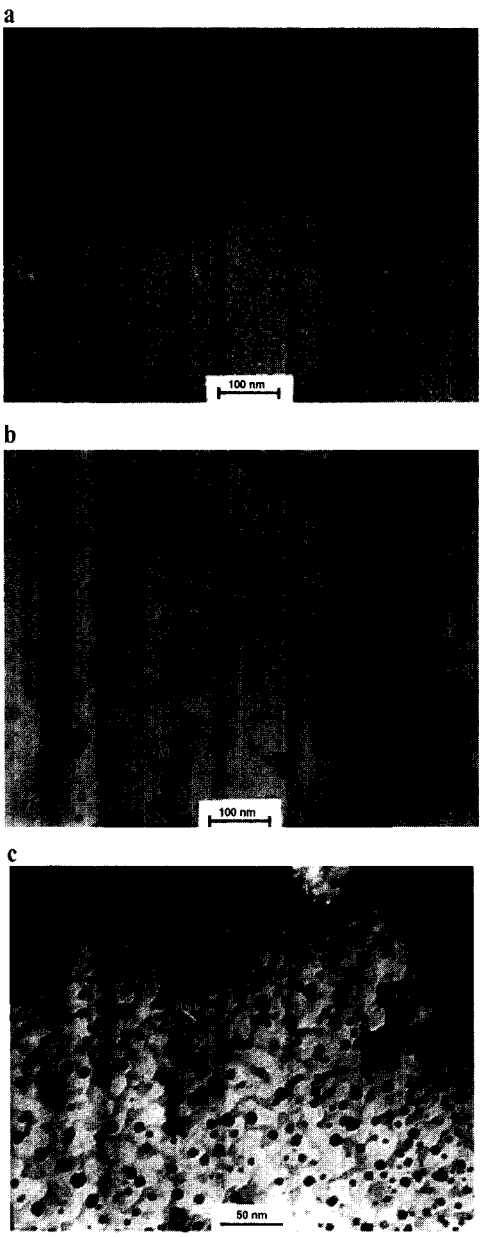


Fig. 3 Transmission electron microscopic images of PI-*b*-PCEMA-*b*-P*t*BA nanospheres at each stage in the synthesis: (a) after PCEMA cross-linking and PI hydroxylation (stained with OsO_4 overnight); (b) after removal of *tert*-butyl groups (stained with OsO_4 over a weekend); and (c) after Fe_2O_3 loading (no staining). (From Ref. [13] Copyright 2000 by the American Chemical Society.)

Fig. 4 Transmission electron microscopic image of Pt nanoparticles prepared in PB-*b*-PEO micelles by $NaBH_4$ reduction. (From Ref. [15] Copyright 1999 by the American Chemical Society.)

As seen from the above examples, many amphiphilic block copolymers form micelles with a functionalized core in the organic medium. When aqueous solutions are preferred, the choice of block copolymers is very limited and metal particle formation is normally more complicated as the pH of the medium should be taken into consideration. A few examples of such block copolymers include P2VP-*b*-PEO and PB-*b*-PEO,[14,15] yet the former block copolymer micellization depends on the pH value:[16] At pH below 5, P2VP-*b*-PEO becomes molecularly soluble in water. At the same time, decrease of pH of the P2VP-*b*-PEO micellar solution after incorporation of metal compounds or metal nanoparticle formation results in no micelle decomposition although the micelle density decreases. In the case of PB-*b*-PEO, micelles formed in water are very dense, so they successfully fulfill two roles: They serve as nanoreactors for Pd, Pt, and Rh nanoparticle formation (Fig. 4) and as metal-particle-containing templates for mesoporous silica casting.[15]

If the P2VP block is a middle block in PS-*b*-P2VP-*b*-PEO triblock copolymer, the "layered," well-defined micelles are formed in water with the PS core, P2VP shell, and PEO corona.[17] Here the P2VP shell serves as a nanoreactor for gold nanoparticle formation. As the shell is formed by the pH-sensitive P2VP block, the authors believe that this system can be useful for encapsulation and/or release of active species. However, one should remember that after metal particle formation, this block loses its ability to dissolve at low pH.[14] So this property can be hardly realized in this system if nanoparticles or metal complexes are formed in the P2VP shell.

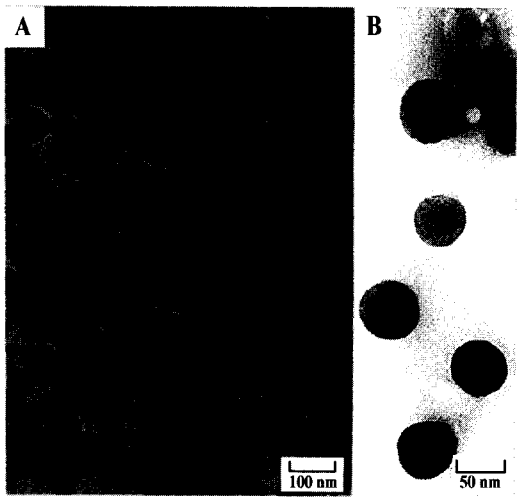


Fig. 5 Transmission electron microscopic images of spherical aggregates in LCM, with Pd/Pt shadowing (A) and without shadowing (B). The dark particles inside the spheres are CdS nanoparticles. (From Ref. [22] Copyright 1998 by the American Chemical Society.)

The formation of spherical assemblies of CdS-containing block copolymer reverse micelles in aqueous solution was reported in Ref. [18]. These stable assemblies were formed by slow addition of water to mixtures of the reverse micelles formed by PS-*b*-PAA and single PS-*b*-PAA chains. Large compound micelles (LCMs) with quantum-confined CdS nanoparticles dispersed throughout a spherical PS stabilized in water by a layer of solubilized hydrophilic chain matrix were obtained. The size of the CdS particles (approximately 3 nm) is determined by the ionic block length of the block copolymer forming the reverse micelle (Fig. 5). The formation of LCMs was found to depend on the amount of the added stabilizing copolymer. This method allows transferring the CdS nanoparticles formed in the micelle cores in organic medium to aqueous medium without loss of stability and nanoparticle aggregation.

Block copolymer micelle coronas

Nanoparticles can be synthesized in the corona of amphiphilic block copolymer micelles. However, if the corona is functionalized, addition of a metal salt can result in immediate formation of large aggregates because of the interaction between micelles and their precipitation; thus this method can be used only in very dilute solutions. If the corona does not contain groups able to coordinate with metal compounds, particle stabilization can be ensured because of the hydrophobic interactions with the hydrophobic core. This feature was used when synthesis of Pd, Pt, Ag, and Au nanoparticles was performed in aqueous solutions of PS-*b*-PEO and PS-*b*-PMAA by reduction of the corresponding salts in block copolymer solutions.[19,20] However, the stability of such systems, solely provided by the hydrophobic interactions with the PS core, is not satisfactory. On the other hand, accessibility of particles in the micelle coronas can be favorable from the viewpoint of catalytic applications.

Enhanced stabilization in the micelle coronas was achieved when hybrid micelles consisting of PS-*b*-PEO and surfactants were formed.[21–23] Surfactant hydrophobic tails were expected to penetrate the PS core while surfactant head groups are located on the micelle core surface or in its vicinity. As shown in Fig. 6, exchange of surfactant counterions for ions of interest would lead to saturation of the core with the given ions. Dynamic light scattering (DLS) and sedimentation in an ultracentrifuge showed that incorporation of positively or negatively charged surfactants results in increase of size and weight of micelles and micellar clusters up to a certain surfactant concentration (which is different for different surfactants). Further increase of surfactant loading (as a rule, above critical micelle concentration for surfactants) results in a moderate decrease of micelle size and weight. Incorporation of surfactant was found to increase the mobility of the PS core and to decrease the mobility of surfactant tails. Both these facts proved comicellization of block copolymer molecules and cationic or anionic surfactants. Ion exchange of surfactant counterions in the PS-*b*-PEO/CPC (cetyl pyridinium chloride) system for $PtCl_6^{2-}$ or $PdCl_4^{2-}$ ions results in saturation of micellar structures with Pt or Pd ions. Subsequent reduction of metal-containing hybrid micellar systems PS-*b*-PEO/CPC/MX_n with $NaBH_4$ or H_2 leads to the formation of metal nanoparticles mainly located within the micelles. The morphology and stability of Pd and Pt nanoparticles synthesized in these systems depends on the metal compound loading and the type of a reducing agent. $NaBH_4$ reduction leads to decomposition

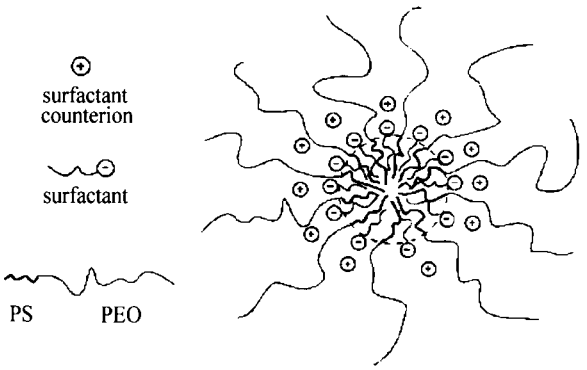


Fig. 6 Schematic image of the PS-*b*-PEO/SDS micelle. (From Ref. [22] Copyright 2000 by Academic Press.)

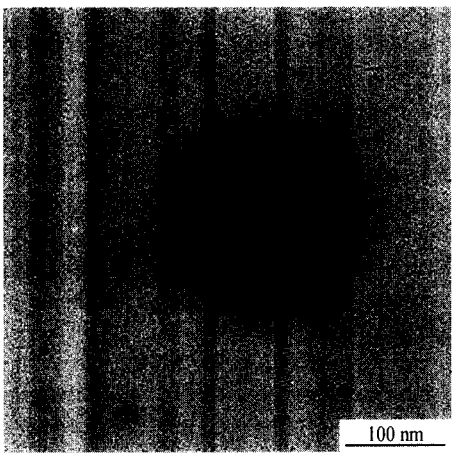


Fig. 7 Transmission electron micrographs of Rh nanoparticles formed in the PS-*b*-PEO/SDS system. (From Ref. [22]. Copyright 2000 by Academic Press.)

of micellar clusters and formation of micelles with embedded nanoparticles. These systems display exceptional stability (for years) if metal salt loading does not exceed 1.24×10^{-2} M. Hydrogen reduction results in metal nanoparticle formation both in micelles and micellar clusters (micelle aggregates), so stability of colloidal solutions is ensured at metal salt concentration of less than 3.36×10^{-3} M. Rh nanoparticles with diameters of 2–3 nm have been obtained in the hybrid micelles formed by PS-*b*-PEO anionic surfactants: sodium dodecylsulfate (SDS) or sodium dodecylbenzosulfonate (SDBS) using Rh cations $[Rh(Py)_4Cl_2]^+$ (Fig. 7). As found, nanoparticle size does not depend on the type of reducing agent (contrary to the nanoparticles formed in other block copolymer solutions),[4,14] but depends on the type of metal.[21–23] This could be governed by the strong interaction of surfactant head groups with growing nanoparticles.

Thus incorporation of surfactants in the block copolymer micelles containing no functional groups allows reliable stabilization of metal nanoparticles of 2–6 nm in size. Using both cationic and anionic surfactants allows one to explore an infinite variety of metal ions and to prepare different kinds of nanoparticles. The disadvantage of these systems is a lack of the direct methods to tune the particle size.

Nanoparticle Formation in Dendrimers

The most commonly used and studied dendrimers (commercially available) are poly(amidoamines) (PAMAM), whose structure is presented in Ref. [24]. By varying the dendrimer size (dendrimer generation) and metal compound loading, one can vary the size of nanoparticles. Dendrimers were used for stabilization of gold and silver nanoparticles with subsequent self-assembling of metal-containing dendrimers on the surface with monolayer formation.[25] Poly(amidoamine) dendrimers with terminal hydroxyl groups were used as templates and stabilizers for controlled synthesis of monodisperse, catalytically active nanoparticles.[26] In the first step, metal ions were absorbed by the dendrimer as a result of the formation of metal complexes with amino groups. Subsequent reduction results in the formation of metal nanoparticles encapsulated in the dendrimers. For hydroxyl-terminated dendrimers, a deficiency of metal exchange between dendrimer molecules provides very narrow particle size distribution and subtle control over nanoparticle growth: The nanoparticle size is exactly determined by the amount of metal atoms loaded in the dendrimer. These metal-containing dendrimers are water-soluble and very stable (no precipitation is observed for months). It is considered that a metal particle is formed in the cavity of the dendrimer. This location provides a certain stabilization of the particles. At the same time, the presence of functional (amino) groups allows additional stabilization of nanoparticles. Metal particle formation resembles the growth of nanoparticles in the cross-linked cores of block copolymer micelles when exchange between micelles does not occur. If PAMAM contains surface (terminal) amino groups, exchange between dendrimer molecules

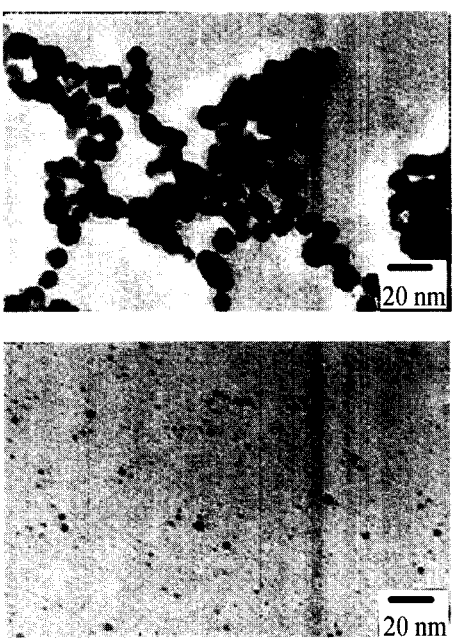


Fig. 8 Electron micrographs of gold colloids and histograms of particle size distribution at molar ratio of surface amino group of G5 and $HAuCl_4$: 1:1 (top) and 4:1 (bottom). (From Ref. [27] Copyright 1998 by the American Chemical Society.)

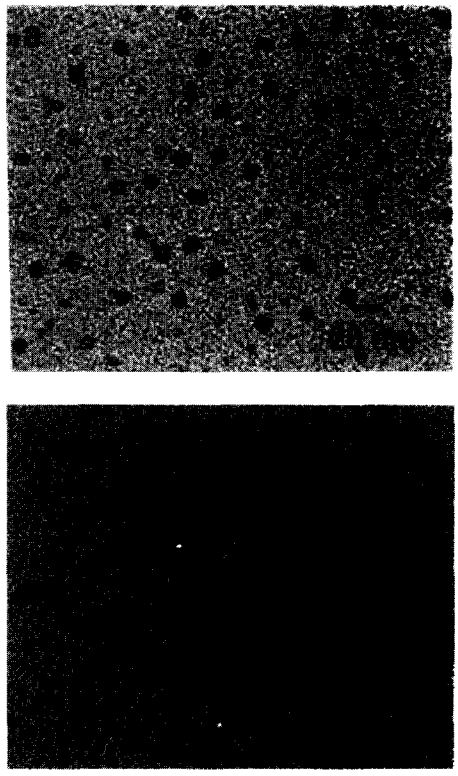


Fig. 9 Transmission electron micrograph of gold containing G8 PAMAM (a) and G10 PAMAM (b) dendrimers obtained for 1:1 loading and slow reduction. In both cases, the dendrimers have been stained with phosphotungstic acid. (From Ref. [29] Copyright 2000 by the American Chemical Society.)

results in a broader particle size distribution than for PAMAM with terminal hydroxyl groups (Fig. 8).[27]

Dendrimer generation is an important factor to influence the particle characteristics.[28] Gold colloid formation upon reduction of a gold salt precursor in protonated PAMAM was studied to follow the influence of reaction conditions and dendrimer generation on the resulting polymer nanocomposite colloids.[29] Methods such as TEM, SANS, and SAXS show that the gold particles are formed inside the dendrimer and located offset from the center (probably in a dendrimer cavity). Lower generation dendrimers aggregate when nanoparticles are formed. Dendrimers of generation 6–9 can template one gold nanoparticle per dendrimer molecule; the number of gold atoms added per dendrimer determines the particle size. These data well agree with those described in Ref. [26]. For generation 10, multiple smaller gold particles per dendrimer were observed and dendrimers aggregated (Fig. 9). Poly(propyleneimine) dendrimers with stearyl end groups, combining both hydrophilic and hydrophobic moieties, were also used for metal particle formation.[30] These dendrimers form inverse micelles in toluene. While the initial dendrimers have a spherical structure with a collapsed core, solubilization of metal salt hydrate leads to the formation of cylindrical multidendrimer structures with swollen, metal-salt-filled dendrimer cores. When the gold salt inside the dendrimers is reduced to form colloidal particles, the cylindrical structure breaks up and spherical nanoparticles are formed. In so doing, the particle sizes are larger than would be expected if the gold-salt loading of one dendrimer formed one particle, indicating that the ions from several dendrimers are combined. Apparently, dendrimers of this kind provide no control over nanoparticle growth. In addition, these dendrimer aggregates, although resembling block copolymer micelles, are less defined and more complex. So advantages of such amphiphilic dendrimers are not evident.

Along with metal particles, a number of semiconductor particles, CdS,[31–33] CdSe,[34] and complex core–shell CdSe/ZnS,[32] were successfully prepared in dendrimers. The absorption and emission of the CdS/PAMAM (hydroxylated) systems are a function of the generation of the dendrimer that is related to the dependence of nanoparticle size on dendrimer generation and the dependence of optical properties on a semiconductor nanoparticle size.[32]

Aqueous assemblies of adamantyl-derivatized poly(propylene imine) (PPI) dendrimers and β-cyclodextrin

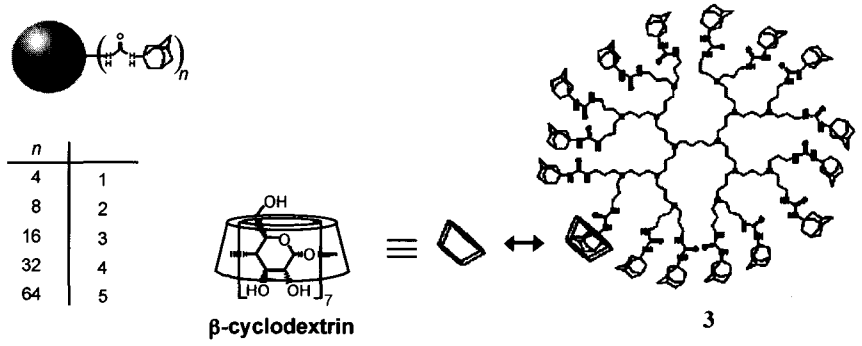


Fig. 10 Poly(propylene imine) dendrimer generations 1–5, solubilized by β-cyclodextrin. (From Ref. [35].)

(β-CD) (Fig. 10) have been used as nanoreactors in the preparation of gold and platinum nanoparticles in water.[35] These particles have been formed by the reduction of aurate or platinate anions in the presence of the generation 4 (**4**_(β-CD)**32**) and 5 (**5**_(β-CD)**40**) assemblies (Fig. 11). Lower generation assemblies did not provide stable nanoparticles. The authors believe that the persistent shape of the adamantyl-derivatized dendrimers and the dense shell of adamantyl–β-CD complexes provide a kinetic barrier for nanoparticle escape, thus prolonging their lifetime. However, particle size distribution inside β-CD-modified dendrimers is not particularly narrow,[35] thus demonstrating poorer control over particle size than in hydroxy-terminated PAMAM.

Nanoparticles in Polyelectrolyte Microgels

Microgels are gel spherical particles having diameters in a nanometer range.[36,37] Internal cross-linking leads to stability of their sizes and properties, while their size (in nanometer range) ensures formation of colloidal solutions. When these microgels are formed by polyelectrolytes,

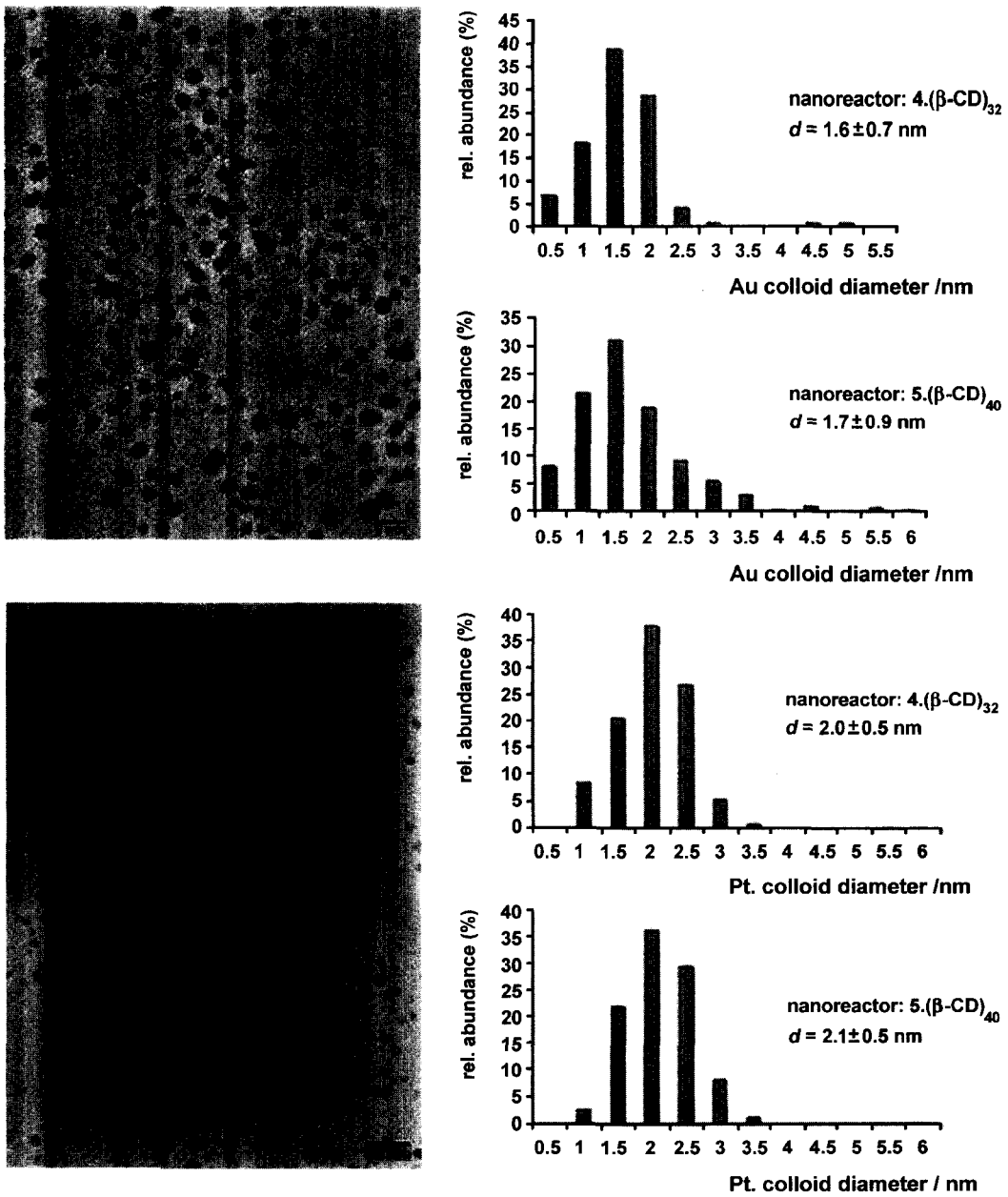


Fig. 11 Transmission electron microscopic images of gold (top) and platinum (bottom) nanoparticles stabilized by the **5**·$(β\text{-CD})_{40}$ assembly and size distributions of colloids stabilized by **4**·$(β\text{-CD})_{32}$ and **5**·$(β\text{-CD})_{40}$ (amine/metal=2:1, H_2O, T=25°C). (From Ref. [35].)

they contain charged groups allowing ion exchange and solubilization in water. This allows considering microgels as nanoreactors for controlling nanoparticle growth. Gold nanoparticle formation was studied in microgels based on sulfonated PS.[38] Morphologies of nanoparticles formed are strongly determined by the degree of microgel cross-linking. The higher the cross-linking density, the higher the probability of the formation of spherical particles embedded in the microgels. The other key factor is the type of reducing agent. Fast reduction ($NaBH_4$) in water leads to gold nanoparticles of 4.5 nm in diameter located in microgels. If $NaBH_4$ is added in alkaline solution (0.1 N NaOH), it slows the reduction and results in 7-nm nanoparticles forming long "threads." Yet only 20% of microgels contain nanoparticles. Thus slow nucleation allows Au clusters and ions to migrate out of microgel areas where aggregation easily occurs. High-resolution TEM allows determining that particles grow from one nucleus and the microgel environment governs the particle shape. In a similar fashion, Pd and Pt nanoparticles have been grown in microgels. These metal-particle-containing microgels were also used as cotemplates (along with amphiphilic PS-*b*-PEO block copolymers) for mesoporous silica formation.[39] Here microgels play a dual role: They are nanoreactors for metal particle formation and pore-forming templates when mesoporous materials are formed.

Nanoparticle Formation in Functionalized Polysilsesquioxane Colloids

Synthesis of functionalized polysilsesquioxane based on hydrolytic condensation of functionalized silanes was recently described in Ref. [40]. Using *N*-(6-aminohexyl)aminopropyltrimethoxysilane (AHAPS) as a precursor, well-defined colloidal particles composed of a nearly fully condensed poly(aminohexyl)(aminopropyl)silsesquioxane (PAHAPS) were synthesized. According to solid state ^{29}Si CP-MAS (cross-polarization magic angle spinning) nuclear magnetic resonance (NMR), PAHAPS structure contains mainly C-$SiO_{3/2}$ species. The sizes of PAHAPS colloids vary in the range 10–200 nm and depend on the reaction conditions: pH, solvent, and AHAPS concentration (Fig. 12). When hydrolytic condensation is carried out in water with no HCl added, self-assembling of AHAPS tails results in the formation of lamellar ordering with Bragg spacing of about 3.0 nm, which matches to two layers of fully extended AHAPS tails. When PAHAPS is fully or partially protonated (HCl is added to water), no ordering occurs. Similarly disordered structures are formed in THF, which is a good solvent for AHAPS tails. Unlike water, THF also facilitates cross-linking between colloids, so the colloids are attached to each other. Interaction of PAHAPS colloids with Pt and Pd salts followed by chemical reduction results in the formation of discrete metal nanoparticles stabilized within the colloids even at metal content of 30–45 wt.% (Fig. 13). The particle size depends on the type of metal compound and, in some cases, on type of the reducing agent. Particles formed after reduction of K_2PtCl_4 within PAHAPS measure ca. 1–2 nm in diameter when the fast reducing agent ($NaBH_4$) is reduced. With the sluggish reducing agent (hydrazine-hydrate), both small particles of 1.5 nm and larger particles with diameters of 2.5–4.0 nm are formed. When metal precursor is K_2PtCl_6, very narrowly distributed particles measuring about 1.5 nm are formed for both types of reducing agents. Apparently, particle

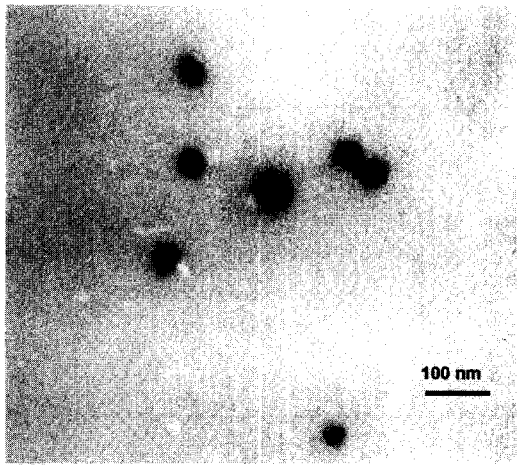


Fig. 12 Transmission electron microscopic image of PAHAPS colloids obtained in water at a precursor concentration of 17 wt.%.

Fig. 13 Transmission electron microscopic image of Pt nanoparticles formed in protonated PAHAPS prepared in water at a concentration of 1.5 wt.%.

nucleation does not influence the particle size when mass transfer of the $PtCl_6^{2-}$ ions is restricted. Poly(aminohexyl)(aminopropyl)silsesquioxane colloids form very strong elastic films because of the intercolloid interactions that makes possible the formation of freestanding films and coatings of different thicknesses on various supports.

Another interesting feature of these particles is that dendrites differing in size and shape are observed to form from aqueous solutions containing PAHAPS colloids loaded with metal salts or metal nanoparticles (Fig. 14).[40] While formation of dendrites is common for a number of salts or metals,[41,42] it has not been reported so far for colloidal particles or nanoparticle-loaded colloidal systems. Because the size and shape of the dendrites produced here can be easily controlled, metal-loaded PAHAPS may be suggested as catalytically active membranes in which select surface coverage is important. Another possible application for PAHAPS loaded with metal nanoparticles is as a conductive layer between two surfaces, as described elsewhere[43] for pure Pd dendrite crystals. At the same time, unlike pure metals, metal-loaded PAHAPS possesses a number of practical advantages, including lower cost and greater structural stability.

Another precursor of interest for synthesis of polymer colloids using sol–gel reaction is octadecyldimethyl(3-trimethoxysilylpropyl)ammonium chloride (ODMACl). The hydrolytic condensation of this precursor both in acidic and basic solutions results in the formation of colloidal particles showing lamellar ordering with Bragg spacing d of about 3.6 nm. This Bragg spacing is significantly smaller than two layers of extended ODMACl tails, so one can assume either tilting of ODMACl chains (they are not perpendicular to the lamellar surface) or interpenetration of the tails of the two layers (or both). As the PODMACl colloids are much larger than twice the d spacing, they should have a multilamellar structure. Because PODMACl colloids contain ionic groups in their tails, they can be subjected to ion exchange that results in replacement of Cl^- ions for ions of interest: $PdCl_4^{2-}$, $PtCl_6^{2-}$, $AuCl_4^-$. The subsequent reduction leads to metal nanoparticle formation within PODMACl colloids. Moreover, independently of reducing agent type, lamellar ordering is preserved, while particle size and shape strongly depend on the nucleation rate. With $NaBH_4$, narrowly distributed spherical particles with a mean diameter of $\sim$2 nm are obtained. When nucleation is slow, particle growth is directed by the lamellar ordering, so rod-like particles of 2×13 nm are formed and their positioning is well regulated by the ordered structure. This suggests new opportunities for growing rod-like particles derived from different metal or semiconductors within ordered polymeric colloids. Such materials can be promising for tailoring optical, magnetic, and electrical properties of nanocomposites by tuning the particle shape.

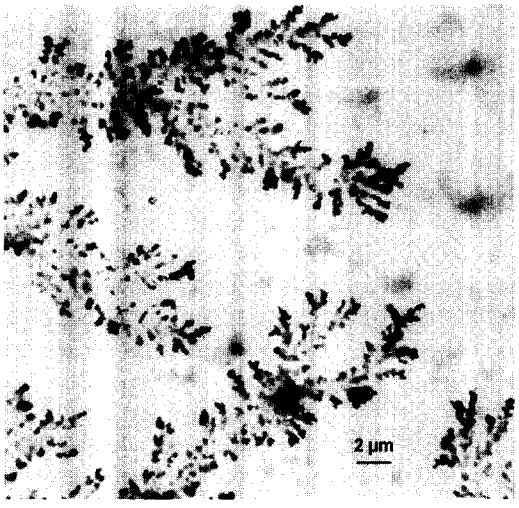


Fig. 14 Transmission electron microscopic image of dendrites formed by PAHAPS colloids filled with Pd nanoparticles.

NANOPARTICLES FORMED ON THE POLYMER COLLOID SURFACE

Micrometer and submicrometer Au-shell PS latex beads have been prepared by combining the self-assembly and seeding methods.[44] The PS beads are first covered with positively charged poly(ethyleneimine) (PEI) via electrostatic interaction and hydrogen bonding; then, the Au shell is formed via the reaction of PEI–PS with NH_2OH and $HAuCl_4$. This method allows efficient control of the gold coverage and leads to relatively stable products. Using TEM, formation of Au clusters and gold nanoparticles on the surface of the PS beads was confirmed. The results suggest that the PEI polymer chains are probably stretched out on the surface of the PS beads. The UV–visible extinction spectrum for gold-covered particles is significantly affected by the plasmon resonance absorption.

For preparation of Au-shell magnetic particles, sulfonated polystyrene beads were first exchanged with Fe^{2+} ions under N_2 gas flow and then treated with NaOH solution. The last stage involves heating the beads at 100 °C for 1 h. Iron oxide layer was later covered with a gold shell as described for Au shell/PS colloids. Self-assembled structures of Au-covered particles were induced by an external magnetic field.[44] This conductive grid was suggested for use as a tunable polarizer for microwaves and milliwaves, as the conducting lines that can be switched into a nonconductive state by an external field. Waveguides for plasmon, i.e., excitation of the plasmon resonance at one end of the lines and detection at the other end, is another interesting application.

Well-dispersed Ag nanoparticles were formed in situ on the surface of poly(*N*-isopropylacrylamide, PNIPAAm)-coated polystyrene microspheres.[45] The surface-grafted PNIPAAm chains serve both as steric stabilizers to prevent the flocculation of the polystyrene particles and interact with silver salts with subsequent adsorption of the forming Ag nanoparticles onto the surfaces of the microspheres. By varying the concentrations of initiator and silver nitrate, the particle sizes and distributions for both Ag nanoparticles and polystyrene microspheres can be altered. As PNIPAAm is a well-known, temperature-sensitive polymer, the PNIPAAm-protected Pt colloids show unusual temperature dependence of activity in the aqueous hydrogenation of allyl alcohol, which can be explained by different density of PNIPAAm layer at different temperatures. In addition, Pt colloids on the microspheres show higher activity than the commercial Pt/C catalyst and retain high activity on recycling in the same reaction. Polystyrene microspheres with silver nanoparticles can be used as the active substrates for surface-enhanced Raman scattering, chemical, electronic, and optical sensors, and photocatalysts for solar energy conversion.

Another example of PS beads covered with nanoparticles is described in Ref. [46]. Surface-functionalized PS microbeads and nanobeads were prepared by grafting the *p*-acetoxystyrene monomer during the last 30 min of the fabrication of polystyrene bead core by emulsifier-free emulsion polymerization followed by hydrolysis of the acetoxy group by a base. The size of the resulting beads is mostly dictated by the size of the core. Hydroxyl-derivatized polystyrene microspheres have been used for anchoring catalytically active silver and ruthenium nanoparticles. This was performed by adsorption of preformed nanoparticles on the functionalized bead surface. The bead formation, surface functionalization, and coating with metal nanoparticles were studied using scanning electron microscopy (SEM), TEM, energy dispersive X-ray spectrometry (EDS), Fourier-transform IR spectrometry, and Auger analysis.

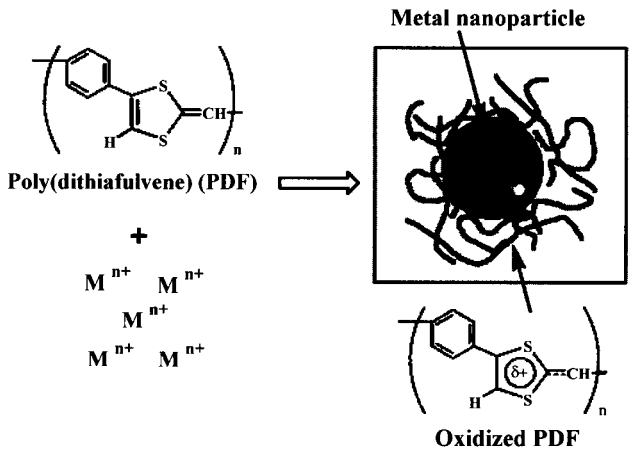


Fig. 15 Schematic illustration of the formation of the PDF-protected metal nanoparticles via reduction of metal ions by the π-conjugated, electron-donating PDF. (From Ref. [47] Copyright 2002 by the American Chemical Society.)

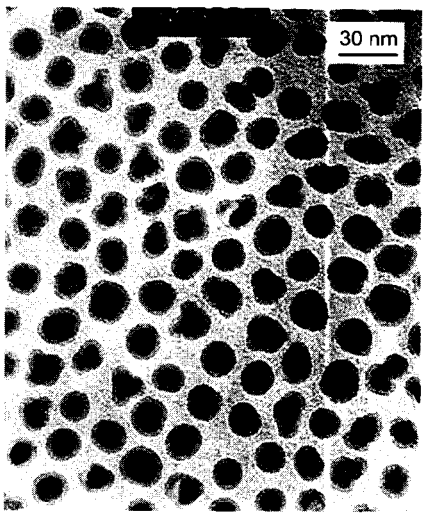


Fig. 16 Transmission electron microscopic image of PIB-coated nanoparticles made with 80 g of $Fe(CO)_5$ and 11 g of PIB–TEPA (tetraethylenepentamine). (From Ref. [48] Copyright 2002 by the American Chemical Society.)

NANOPARTICLES WITH ADSORBED POLYMER LAYER

Any nanoparticle synthesis in polymer solution normally results in the adsorption of polymer molecules on the nanoparticle surface. Usually, these materials are not considered as metalled polymer colloids. Nevertheless, in some cases including ones cited below, adsorption or chemical interaction of polymers with nanoparticles is accompanied with formation of a polymer layer of different nature than that used for nanoparticle synthesis.

Reduction of Pd, Pt, or Au ions in dimethyl sulfoxide (DMSO) solution by a π-conjugated poly(dithiafulvene) (PDF) having strong electron-donating properties resulted in the adsorption of oxidized polymer on the forming nanoparticle surface, which protected the metal nanoparticles (Fig. 15).[47] As found, all the DF units of PDF were uniformly oxidized by electron transfer, which means that they participated in the reduction, and resided on the particle surface. The average size of the formed Pt nanoparticles increased with the increases of the PDF

concentration and reaction temperature. All of the Pt nanoparticles formed at different temperatures showed narrow size distribution, high dispersibility, and great thermal stability.

Magnetic materials comprised of polymer–iron nanoparticle composites have been prepared by thermal decomposition of iron pentacarbonyl in the presence of ammonia and several types of polymeric dispersants.[48] The nanoparticles consist of metallic cores, each of which is coated with a strongly bound polymer layer (Fig. 16). The polyisobutylene- and polyethylene-based dispersants lead to more uniform particle sizes and the materials are composed principally of individual core–shell particles. Large, complex particles, formed by the aggregation of smaller particles, as well as simple core–shell particles, were obtained with polystyrene-based dispersants in 1-methylnaphthalene. Aggregation leading to complex particles was found to be a result of a mismatch between the solubility parameters of the dispersant and the reaction solvent. When the solvent was better matched with the solubility parameter of polystyrene, simple core–shell nanoparticles of fairly uniform size were obtained. Electron diffraction revealed that the core material is iron. High-resolution TEM showed highly ordered or crystalline regions within the polymer shell, possibly because of the dense packing of the strongly bound dispersant chains (Fig. 17). Magnetic interactions between particles cause formation of secondary structures such as clusters, coils, loops, and strings of particles. The nanocomposites can be dispersed in organic solvents or cast as films. The polystyrene–iron composites are rigid solids at room temperature but can be melted at temperatures of 65–100°C. The nanoparticle composites showed different magnetic behavior depending on the particle size. Samples with the smallest particles proved to be superparamagnetic but their saturation magnetizations were low. Hysteresis was observed for materials with larger particles, and the materials possessed larger magnetizations. The magnetization was correlated with the particle size where samples with larger particles showed higher magnetizations. The samples with higher iron contents, achieved using higher iron pentacarbonyl loadings or by removing unbound polymer dispersant, showed the highest magnetization.

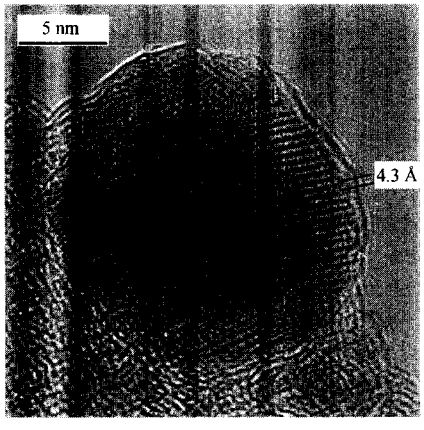


Fig. 17 High-resolution TEM image of PIB-coated nanoparticles made with 60 g of $Fe(CO)_5$ and 11 g of PIB–TEPA. Lattice planes are visible in the shell region of the particles. (From Ref. [48] Copyright 2002 by the American Chemical Society.)

CONCLUSION

This entry clearly shows several approaches to constructing nanocomposite polymer colloids containing nanoparticles. In all cases, polymer colloids are soluble in organic or aqueous media (depending on the exterior of the colloids) and retain solubility after nanoparticle formation. This key feature allows the formation of thin deposited or freestanding films (the latter were obtained with block copolymer micelles and functionalized polysilsesquioxane colloids) that makes possible a number of important applications for nanolithography in optical and magnetic materials. For catalytic applications, both homogeneous (solutions) and heterogeneous systems (after deposition on the support) proved to be promising with the polymer colloids filled with nanoparticles. It is worth mentioning that the polymer nanoenvironment significantly changes the nanoparticle properties via modification of the nanoparticle surface with polymer groups; therefore the choice of polymer type and structural organization of polymer colloids play a crucial role in material properties and should be taken into consideration for specific material application.

REFERENCES

1. Hamley, I.U. *The Physics of Block Copolymers*; Oxford University Press: Oxford, UK, 1998.
2. Förster, S.; Antonietti, M. Amphiphilic block copolymers in structure-controlled nanomaterial hybrids. Adv. Mater. **1998**, *10* (3), 195–217.
3. Antonietti, M.; Henz, S. Supermolecular structures at polymers. A way towards intelligent materials? Nachr. Chem. Tech. Lab. **1992**, *40* (3), 308–314.
4. Antonietti, M.; Wenz, E.; Bronstein, L.; Seregina, M. Synthesis and characterization of noble metal colloids in block copolymer micelles. Adv. Mater. **1995**, *7* (12), 1000–1005.

5. Spatz, J.P.; Roescher, A.; Möller, M. Gold nanoparticles in micellar poly(styrene)-*b*-poly(ethylene oxide) films. Size and interparticle distance control in monoparticulate films. Adv. Mater. **1996**, *8* (4), 337–340.
6. Moffitt, M.; McMahon, L.; Pessel, V.; Eisenberg, A. Size control of nanoparticles in semiconductor-polymer composites. 2. Control via sizes of spherical ionic microdomains in styrene-based diblock ionomers. Chem. Mater. **1995**, *7* (6), 1185–1192.
7. Saito, R.; Okamura, S.; Ishizu, K. Introduction of colloidal silver into poly(2-vinylpyridine) microdomains of microphage-separated poly(styrene-*b*-2-vinylpyridine) film. 3. Poly(2-vinylpyridine) spherical microdomain. Polymer **1993**, *34* (6), 1189–1195.
8. Bronstein, L.; Antonietti, M.; Valetsky, P. Metal Colloids in Block Copolymer Micelles: Formation and Material Properties. In *Nanoparticles and Nanostructured Films*; Fendler, J.H., Ed.; Wiley-VCH Verlag: Weinheim, 1998; 145–171.
9. Seregina, M.V.; Bronstein, L.M.; Platonova, O.A.; Chernyshov, D.M.; Valetsky, P.M.; Hartmann, J.; Wenz, E.; Antonietti, M. Preparation of noble-metal colloids in block copolymer micelles and their catalytic properties in hydrogenation. Chem. Mater. **1997**, *9* (4), 923–931.
10. Bronstein, L.M.; Chernyshov, D.M.; Volkov, I.O.; Ezernitskaya, M.G.; Valetsky, P.M.; Matveeva, V.G.; Sulman, E.M. Structure and properties of bimetallic colloids formed in polystyrene-*block*-poly-4-vinylpyridine micelles: Catalytic behavior in selective hydrogenation of dehydrolinalool. J. Catal. **2000**, *196* (2), 302–314.
11. Platonova, O.A.; Bronstein, L.M.; Solodovnikov, S.P.; Yanovskaya, I.M.; Obolonkova, E.S.; Valetsky, P.M.; Wenz, E.; Antonietti, M. Cobalt nanoparticles in block copolymer micelles. Preparation and properties. Colloid Polym. Sci. **1997**, *275* (5), 426–431.
12. Rutnakornpituk, M.; Thompson, M.S.; Harris, L.A.; Farmer, K.E.; Esker, A.R.; Riffle, J.S.; Connolly, J.; St. Pierre, T.G. Formation of cobalt nanoparticle dispersions in the presence of polysiloxane block copolymers. Polymer **2002**, *43* (8), 2337–2348.
13. Underhill, R.S.; Liu, G. Triblock nanospheres and their use as templates for inorganic nanoparticle preparation. Chem. Mater. **2000**, *12* (8), 2082–2091.
14. Bronstein, L.M.; Sidorov, S.N.; Valetsky, P.M.; Hartmann, J.; Coelfen, H.; Antonietti, M. Induced micellization by interaction of poly(2-vinylpyridine)-*block*-poly(ethylene oxide) with metal compounds. Micelle characteristics and metal nanoparticle formation. Langmuir **1999**, *15* (19), 6256–6262.
15. Bronstein, L.; Kraemer, E.; Berton, B.; Burger, C.; Foerster, S.; Antonietti, M. Successive use of amphiphilic block copolymers as nanoreactors and templates: Preparation of porous silica with metal nanoparticles. Chem. Mater. **1999**, *11* (6), 1402–1405.
16. Martin, T.J.; Prochazka, K.; Munk, P.; Webber, S.E. pH-dependent micellization of poly(2-vinylpyridine)-*block*-poly(ethylene oxide). Macromolecules **1996**, *29* (18), 6071–6073.
17. Gohy, J.-F.; Willet, N.; Varshney, S.; Zhang, J.-X.; Jerome, R. Core–shell–corona micelles with a responsive shell. Angew. Chem., Int. Ed. **2001**, *40* (17), 3214–3216.
18. Moffitt, M.; Vali, H.; Eisenberg, A. Spherical assemblies of semiconductor nanoparticles in water-soluble block copolymer aggregates. Chem. Mater. **1998**, *10* (4), 1021–1028.
19. Mayer, A.B.R.; Mark, J.E.; Morris, R.E. Palladium and platinum nanocatalysts protected by amphiphilic block copolymers. Polym. J. **1998**, *30* (3), 197–205.
20. Mayer, A.B.R.; Mark, J.E. Transition metal nanoparticles protected by amphiphilic block copolymers as tailored catalyst systems. Colloid Polym. Sci. **1997**, *275* (4), 333–340.
21. Bronstein, L.M.; Chernyshov, D.M.; Timofeeva, G.I.; Dubrovina, L.V.; Valetsky, P.M.; Obolonkova, E.S.; Khokhlov, A.R. Interaction of polystyrene-*block*-poly(ethylene oxide) micelles with cationic surfactant in aqueous solutions. Metal colloid formation in hybrid systems. Langmuir **2000**, *16* (8), 3626–3632.
22. Bronstein, L.M.; Chernyshov, D.M.; Timofeeva, G.I.; Dubrovina, L.V.; Valetsky, P.M.; Khokhlov, A.R. The hybrids of polystyrene-*block*-poly(ethylene oxide) micelles and sodium dodecyl sulfate in aqueous solutions: Interaction with Rh ions and Rh nanoparticle formation. J. Colloid Interface Sci. **2000**, *230* (1), 140–149.
23. Bronstein, L.M.; Chernyshov, D.M.; Vorontsov, E.; Timofeeva, G.I.; Dubrovina, L.V.; Valetsky, P.M.; Kazakov, S.; Khokhlov, A.R. Comicellization of polystyrene-*block*-poly(ethylene oxide) with cationic and anionic surfactants in aqueous solutions: Indications and limits. J. Phys. Chem., B **2001**, *105* (38), 9077–9082.
24. Zeng, F.; Zimmerman, S.C. Dendrimers in supramolecular chemistry: From molecular recognition to self-assembly. Chem. Rev. **1997**, *97* (5), 1681–1712.
25. Bar, G.; Rubin, S.; Cutts, R.W.; Taylor, T.N.; Zawodzinski, T.A., Jr. Dendrimer-modified silicon oxide surfaces as platforms for the deposition of gold and silver colloid monolayers: Preparation

method, characterization, and correlation between microstructure and optical properties. Langmuir **1996**, *12* (5), 1172.
26. Zhao, M.; Crooks, R.M. Dendrimer-encapsulated Pt nanoparticles: Synthesis, characterization, and applications to catalysis. Adv. Mater. **1999**, *11* (3), 217–220.
27. Esumi, K.; Suzuki, A.; Aihara, N.; Usui, K.; Torigoe, K. Preparation of gold colloids with UV irradiation using dendrimers as stabilizer. Langmuir **1998**, *14* (12), 3157–3159.
28. Crooks, R.M.; Lemon, B.I.; Sun, L.; Yeung, L.K.; Zhao, M. Dendrimer-encapsulated metals and semiconductors: Synthesis, characterization, and applications. Topics Curr. Chem. **2001**, *212* (Dendrimers III), 81–135.
29. Gröhn, F.; Bauer, B.J.; Akpalu, Y.A.; Jackson, C.L.; Amis, E.J. Dendrimer templates for the formation of gold nanoclusters. Macromolecules **2000**, *33* (16), 6042–6050.
30. Gröhn, F.; Bauer, B.J.; Amis, E.J. Hydrophobically modified dendrimers as inverse micelles: Formation of cylindrical multidendrimer nanostructures. Macromolecules **2001**, *34* (19), 6701–6707.
31. Sooklal, K.; Huang, J.; Murphy, C.J.; Hanus, L.; Ploehn, H.J. Inorganic quantum dot-organic dendrimer nanocomposite materials. Mater. Res. Soc. Symp. Proc. **1999**, *576*, 439–444.
32. Lemon, B.I.; Crooks, R.M. Preparation and characterization of dendrimer-encapsulated CdS semiconductor quantum dots. J. Am. Chem. Soc. **2000**, *122* (51), 12886–12887.
33. Donners, J.J.J.M.; Hoogenboom, R.; Schenning, A.P.H.J.; van Hal, P.A.; Nolte, R.J.M.; Meijer, E.W.; Sommerdijk, N.A.J.M. Fabrication of organic–inorganic semiconductor composites utilizing the different aggregation states of a single amphiphilic dendrimer. Langmuir **2002**, *18* (7), 2571–2576.
34. Seker, F.; Ellis, A.B. Probing polymer adsorption using an emissive semiconductor substrate: Adsorption of poly(acrylic acid) onto cadmium selenide. Macromolecules **2000**, *33* (2), 582–589.
35. Michels, J.J.; Huskens, J.; Reinhoudt, D.N. Dendrimer–cyclodextrin assemblies as stabilizers for gold and platinum nanoparticles. J. Chem. Soc., Perkin Trans. **2002**, *2* (1), 102–105.
36. Antonietti, M. Microgels—Polymers with a special molecular architecture. Angew. Chem. **1988**, *100* (12), 1813–1817.
37. Antonietti, M.; Bremser, W.; Schmidt, M. Microgels: Model polymers for the crosslinked state. Macromolecules **1990**, *23* (16), 3796–3805.
38. Antonietti, M.; Gröhn, F.; Hartmann, J.; Bronstein, L. Nonclassical shapes of noble-metal colloids by synthesis in microgel nanoreactors. Angew. Chem., Int. Ed. **1997**, *36* (19), 2080–2083.
39. Whilton, N.T.; Berton, B.; Bronstein, L.; Hentze, H.-P.; Antonietti, M. Organized functionalization of mesoporous silica supports using prefabricated metal–polymer modules. Adv. Mater. **1999**, *11* (12), 1014–1018.
40. Bronstein, L.M.; Linton, C.; Karlinsey, R.; Stein, B.; Svergun, D.I.; Zwanziger, J.W.; Spontak, R.J. Synthesis of metal-loaded poly(aminohexyl)(aminopropyl)silsesquioxane colloids and their self-organization into dendrites. Nano Lett. **2002**, *2* (8), 873–876.
41. Honjo, H.; Ohta, S. Dense-branching morphology of an ammonium chloride crystal. Phys. Rev., A **1992**, *45* (12), R8332–R8335.
42. Zeng, H.C.; Lim, L.C. Secondary ionic forces in lead molybdate melt solidification. J. Mater. Res. **1998**, *13* (6), 1426–1429.
43. Bernier, W.E.; Bundga, E.G. Conductive adhesive having a palladium matrix interface between two metal surfaces. **2001**. US 6331119.
44. Ji, T.; Lirtsman, V.G.; Avny, Y.; Davidov, D. Preparation, characterization, and application of Au-shell/polystyrene beads and Au-shell/magnetic beads. Adv. Mater. **2001**, *13* (16), 1253–1256.
45. Chen, B.C.-W.; Chen, M.-Q.; Serizawa, T.; Akashi, M. In-situ formation of silver nanoparticles on poly(*N*-isopropylacrylamide)-coated polystyrene microspheres. Adv. Mater. **1998**, *10* (14), 1122–1126.
46. Greci, M.T.; Pathak, S.; Mercado, K.; Prakash, G.K.S.; Thompson, M.E.; Olah, G.A. Poly(*p*-hydroxystyrene) grafted polystyrene nanospheres: Excellent hosts for silver and ruthenium nanoparticles. J. Nanosci. Nanotech. **2001**, *1* (1), 3–6.
47. Zhou, Y.; Itoh, H.; Uemura, T.; Naka, K.; Chujo, Y. Synthesis of novel stable nanometer-sized metal (M=Pd, Au, Pt) colloids protected by a pi-conjugated polymer. Langmuir **2002**, *18* (1), 277–283.
48. Burke, N.A.D.; Stoever, H.D.H.; Dawson, F.P. Magnetic nanocomposites: Preparation and characterization of polymer-coated iron nanoparticles. Chem. Mater. **2002**, *14* (11), 4752–4761.

Polymer Nanocomposites with Particle and Carbon Nanotube Fillers

B. J. Ash
Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.

A. Eitan
L. S. Schadler
Rensselaer Polytechnic Institute, Troy, New York, U.S.A.

INTRODUCTION

Polymer nanocomposites are composites with a polymer matrix and a filler with at least one dimension less than 100 nm. The fillers can be plate-like (clays), high aspect ratio nanotubes, and lower aspect ratio or equiaxed nanoparticles. While some nanofilled composites (carbon black[1] and fumed silica[2,3]-filled polymers) have been used for over a century, in recent years the dedicated research and development of nanofilled polymers has greatly increased. This is due to our increased ability to synthesize and manipulate a broad range of nanofillers and significant investment by government and industry in this field.

Current interest in nanocomposites has been generated and maintained because nanoparticle and carbon nanotube-filled polymers exhibit unique combinations of properties not achievable with traditional composites. For example, the inclusion of equiaxed nanoparticles in thermoplastics, and particularly semicrystalline thermoplastics, increases the yield stress, the tensile strength, and Young's modulus[4] compared to pure polymer. Other examples include scratch-resistant transparent amorphous thermoplastic coatings.[5] These combinations of properties can be achieved because of the small size of the fillers, the large surface area the fillers provide, and in many cases the unique properties of the fillers themselves. As will be shown, in many cases these large changes in the material properties require small to modest nanofiller loadings. Unlike traditional micron-filled composites, these novel fillers often alter the properties of the entire polymer matrix while, at the same time, imparting new functionality because of their chemical composition and nanoscale size.

This article will give a general introduction to polymer nanocomposites and address what is unique to nanofillers compared to traditional micron-scale fillers. The second section will briefly address nanofiller surface modification and the third will provide specific examples of mechanical, electrical, and optical properties in nanoparticle-filled polymers. The last section provides a detailed description of the mechanical properties of nanotube-filled polymers and a brief description of some electrical and optical properties that have been reported.

WHAT MAKES NANOCOMPOSITES UNIQUE

The small size of nanofillers leads to several factors that distinguish nanocomposites from traditional composites. First of all, nanofillers are small mechanical, optical, and electrical defects compared to micron-scale fillers. This means that the addition of nanofillers to a polymer does not necessarily lead to a decrease in the ductility of the polymer and in some cases can increase it.[6,7] It also means that below about 50 nm,[8] many fillers do not scatter light significantly. Thus it is possible to make composites with altered electrical or mechanical properties that maintain their optical clarity. Finally, as small electrical defects, nanofillers do not concentrate electromagnetic fields as sharply as micron-scale fillers and indeed may act to trap charge and increase the electrical breakdown strength of polymers.

Secondly, although many properties of a material are said to be intrinsic, they often depend upon matter being assembled above a critical length scale. When the nanoparticles decrease below this size, the properties of the particles can differ significantly from the bulk material; thus variations in melting temperature, color, magnetization, and charge capacity are often observed.[9]

Third, the small size of the fillers leads to an exceptionally large interfacial area in the composites. Fig. 1a shows the surface area per unit volume as a function of particle size for spherical particles that are ideally dispersed. If one compares the surface area of a 10-μm carbon fiber to that of a 1-nm single-walled nanotube (SWNT) for the same total volume of the two, the surface area increases by a factor of 10,000. In addition (Fig. 1b), the interparticle spacing decreases such that at small

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009067

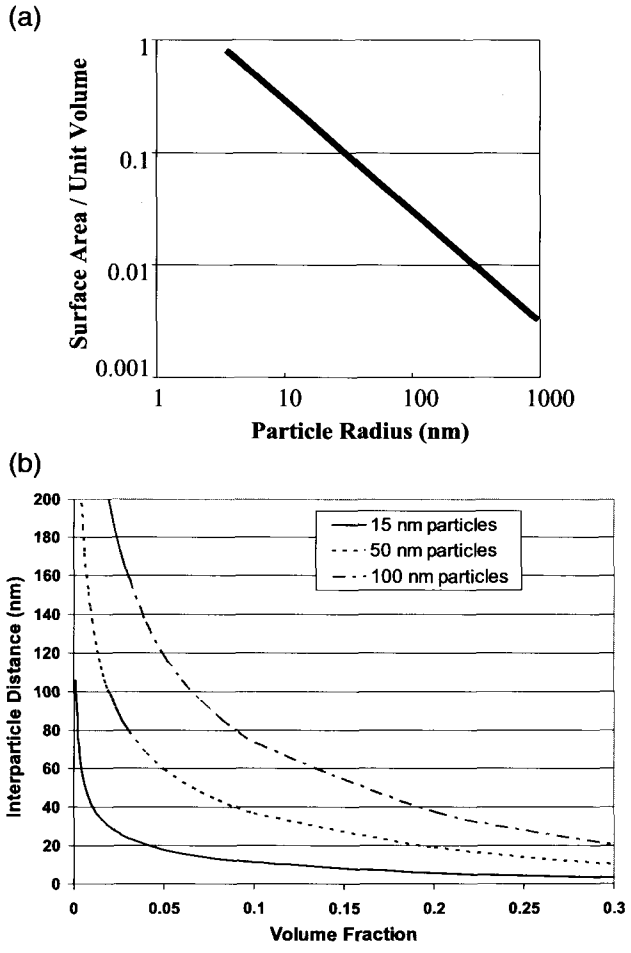


Fig. 1 (a) The surface area per unit volume as a function of particle size for spherical particles showing the large surface area in nanoparticles. (b) The interparticle spacing of nanoparticles arranged on a simple cubic lattice showing the variation with particle size (15, 50, and 100 nm diameter) and volume fraction.

volume fractions of filler, the interparticle spacing is similar to the radius of gyration of the polymer ($\sim$100 Å). The high surface area becomes even more significant when one considers that there is an interaction zone (IZ) surrounding the filler. This is a region in which the structure and properties have been altered because of the presence of the filler. It could be a region of altered chemistry, polymer conformation, chain mobility, degree of cure, or crystallinity. This zone of affected polymer has been approximated to be between 2 and 9 nm thick,[10] but may be much larger. If we assume that this IZ is about 10 nm in thickness, then at 2.5 vol.% of a 20-nm equiaxed nanoparticle well dispersed, 37% of the polymer has different properties from the bulk polymer. Therefore the IZ can be a significant portion, if not the entire bulk, of the

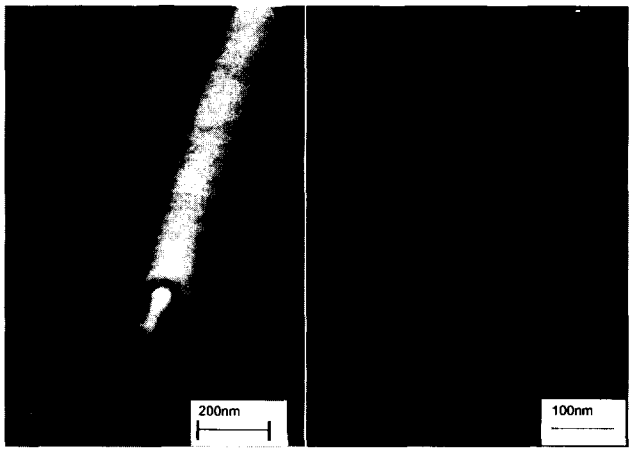


Fig. 2 Polymer-coated nanotubes observed in the fracture surface of MWNT reinforced polycarbonate composite. (From Ref. [16].) (*View this art in color at www.dekker.com.*)

matrix. Thus the nanofillers can alter the expected properties of the composite considerably.

An example of the influence of the IZ on behavior can be seen by monitoring the glass transition temperature, T_g. The T_g of a bulk part can be raised and lowered with the addition of nanoparticles due to the immobilization of polymer chains by the particles or, conversely, an increase in polymer mobility due to noninteracting particles. Both increasing[11–13] and decreasing[14,15] T_g cases have been shown. The physical nature and extent of this IZ has recently been probed through some recent work on multiwalled carbon nanotubes (MWNTs).[16] In this study, a solvent processing method was used to make MWNT polymer nanocomposites. Upon observation of the composite fracture surface, a polymer layer was observed on the nanotubes (Fig. 2) that had pulled out of the opposing

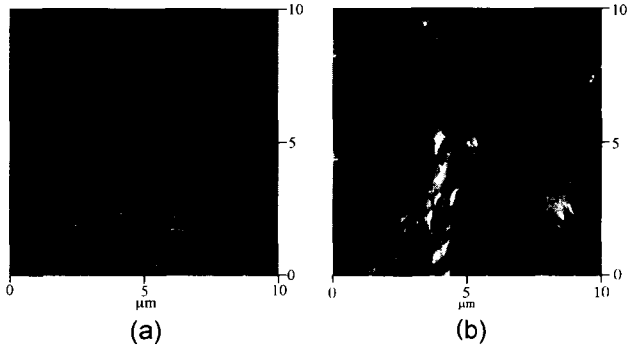


Fig. 3 Atomic force micrographs showing the change in crystalline morphology for unfilled and nanofilled low-density polyethylene, (a) neat-low density polyethylene, (b) low-density polyethylene with 5 wt.% titania nanoparticles. (*View this art in color at www.dekker.com.*) (From Ref. [18].)

side of the fracture. This "sheath" was confirmed to be polymer from the matrix, but with obviously altered thermal and mechanical properties from the bulk. The thickness of this interfacial layer increased with chemical modification of the nanotubes.

Nanoparticles can also influence the polymerization, curing, or crystallization aspects of polymer synthesis. For example, nanoparticles can serve as nucleation sites in semicrystalline polymers and result in changes in crystalline content and spherulite structure.[17] Fig. 3 shows an AFM micrograph of unfilled polyethylene and nanoparticle-filled polyethylene. Note that in the case of the nanofilled polyethylene, the crystalline structure is much less organized.[18] Similar results have been seen in polyethylene terephthalate (PET).[19]

SURFACE MODIFICATION OF NANOFILLERS

The nanoscale sizes and subsequently higher surface energies of nanofillers lead to some unique challenges in the processing of these materials. The most critical of these challenges is dispersion of the nanofiller. Any agglomeration of the filler reduces the interfacial area in the composites and thus reduces the opportunity to take advantage of the unique nanofiller properties mentioned in the previous section. Thus aggregated nanoparticles are simply micron fillers.

Controlling the size and degree of agglomeration of nanoparticles is difficult due to their large radius of curvature and subsequent increase in surface energy. As these high surface area fillers tend to aggregate, silanes and organotitanates are used extensively both to tailor the particle surface properties to mimic the surrounding matrix and to lower their surface energy and reduce their tendency to agglomerate.[20–22] Other methods to alter the surface properties of the nanoparticles include radiation grafting,[23] chemical vapor deposition, and a host of complicated synthesis procedures that attempt to polymerize polymer chains off of initiating agents coupled to the surface.[24–26] An excellent review by Caruso[27] provides an extensive background on the modification of nanoparticle surfaces. In addition to achieving better dispersion, these techniques control the nature of the interaction between the nanofillers and the polymer and thus the properties and size of the IZ.

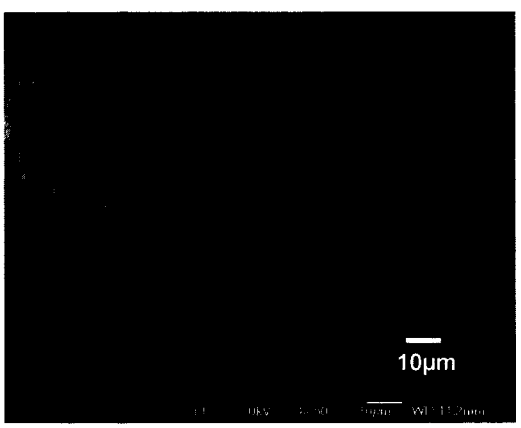


Fig. 4 An SEM micrograph showing the bundled nature of multiwalled carbon nanotubes grown using a chemical vapor deposition process.

Carbon nanotubes not only tend to agglomerate but are often prepared in a bundle-like structure, as illustrated in Fig. 4. In order to take advantage of their high surface area for interacting with the polymer, the bundles have to be separated into individual nanotubes. For MWNT this can usually be accomplished with sonication.[28,29] For SWNT, exfoliation is a more difficult process but progress is being made.[30–34] The challenge is in exfoliating the bundles without shortening the SWNT and introducing significant numbers of defects. Once the nanotubes have been separated, it is important to disperse them uniformly in the polymer matrix, preventing agglomeration of nanotubes. This is accomplished with functional groups which also mediates the interaction of the nanotubes with the polymer.

Noncovalent surface modification of carbon nanotubes includes all treatments that cause a change in the functional groups that face the solvent (or the polymer), without modifying the chemical nature of the nanotube. The advantage of the noncovalent surface modification is that the basic structure and hence the mechanical and electrical properties of the tubes are not affected due to the modification.[35] Noncovalent attachment is possible if there is a secondary bonding between these groups and the surface of the nanotube.[36,37] For example, wrapping of the nanotubes by polymer chains, in particular conjugated polymers, has been observed.[36,38–40]

Covalent attachment of chemical groups to the outer wall of the nanotubes can occur particularly at defect sites. One example is the attack of the defect sites by concentrated nitric acid, in order to form carboxylic acid groups.[41] The reaction with nitric acid also eliminates the catalysts that are left from the nanotubes preparation process. This reaction has been applied on SWNT and MWNT. The resultant carboxylated nanotubes can then be further covalently modified by means of reactions based on the carboxylic acid groups.[42,43] There are also other variations of the chemical oxidation of carbon nanotubes (for example, applying a mixture of sulfuric acid and H_2O_2[44]). The presence of carboxylic acid groups on the nanotube walls enables various reactions for the further attachment of functional groups.[43,45–50]

Other types of covalent surface modification of carbon nanotubes are based on chemical reactions between the

carbon–carbon bond structures and specific reagents such as fluorination[51] or on radical attachments.[52,53] These reactions and many others[54–58] enable tailoring of the nanofillers to the specific application and environment. The potential for using surface modification to improve the properties of nanotube-reinforced polymer composites has just begun to be explored.

NANOPARTICLE-FILLED POLYMER PROPERTIES

Nanoparticle fillers are relatively equiaxed with diameters less than 100 nm (Fig. 5). Nanoparticles (carbon black) have been used as fillers in polymers for the past century, but the advent of new particle synthesis techniques has made available industrial quantities of ceramic, piezoelectric, metal, and even polymer nanoparticles for use in polymer nanocomposite research. Synthesis of these nanoparticles is accomplished using a variety of methods, each with its own advantages and disadvantages.[59] This section focuses mainly on the unique mechanical, thermal, optical, and electrical properties obtained when nanoparticles are used as fillers in polymer systems.

Mechanical Properties

One of the primary reasons for adding fillers to polymers is to improve mechanical performance. Traditional polymer composites filled with micron-size fillers often show improvements in stiffness and heat resistance in the form

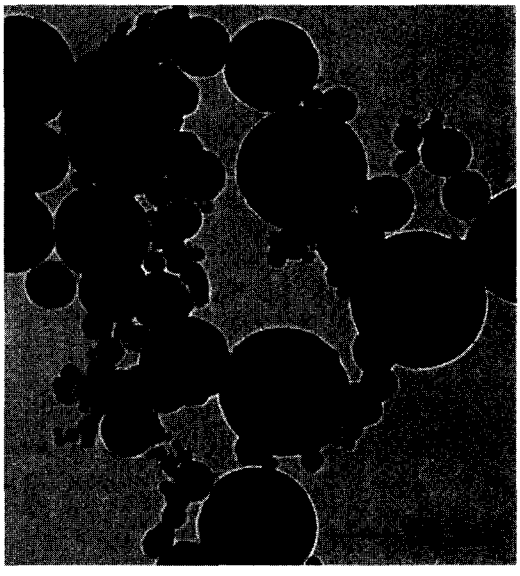


Fig. 5 Typical poly-disperse oxide nanoparticles synthesized in a forced gas condensation process.

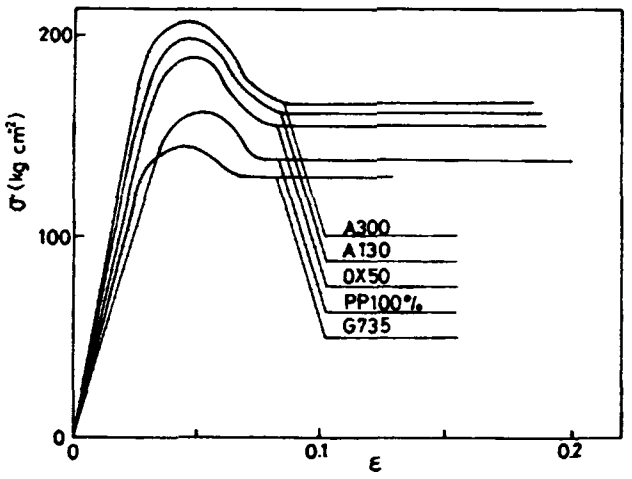


Fig. 6 Stress–strain curves of a silica/polypropylene(PP) nanocomposite with fillers of four different diameters (A300 and A130 = nanosilica-filled PP, PP100% = neat polymer, and G735 = micron-filled PP). (From Ref. [4].)

of increases in modulus, yield strength, and glass transition temperature.[4,60] In micron-filled composites, unfortunately, this often comes at the cost of a substantial reduction in ductility, and sometimes in impact strength, because of stress concentrations caused by the fillers. Well-dispersed nanofillers, on the other hand, are able to improve modulus and strength, and to maintain or even improve ductility because they are much smaller than the critical crack size for polymers and need not initiate failure. In addition, large amounts of traditional fillers are often required to achieve the desired properties, often diminishing the weight-savings gained in using low-density polymers.

Polymer nanocomposites have been shown to provide unique combinations of mechanical and thermal properties often at very low filler weight fractions. For example, dramatic improvements in the yield stress (30%) and Young's modulus (170%) have been shown in nanofilled polypropylene compared to micron-filled polypropylene.[61] These composites also showed no decrease in the strain-to-failure when filled with silica ranging from 7 to 40 nm in diameter (Fig. 6). In addition, nylon 6 filled with 50-nm silica particles displayed increases in tensile strength (15%), strain-to-failure (150%), Young's modulus (23%), and impact strength (78%) with only 5 wt.% nanoparticles.[62] In a rubbery polyurethane elastomer, Petrovic and Zhang[63,64] found that a sixfold increase in the elongation-at-break and a threefold increase in the modulus were achievable with 40 wt.% 12-nm silica compared to micron-size filler that embrittled the polymer. Furthermore, the nanocomposites showed no reduction in transparency even at these relatively high loadings

unlike the micron-size filled systems. In attempting to modify an already two-phase system, Zhang et al.[65] filled high impact polystyrene (HIPS) with nanoparticles and showed increases in notched impact, tensile strength, and elastic modulus of the blend. In a study by Ash et al., the mode of yielding in polymethylmethacrylate, a brittle thermoplastic polymer, was changed from dilatational (craze) to shear by the incorporation of nanoparticles that showed no affinity for the matrix polymer.[66] Extensions as high as 80% but averaging 30% were routinely seen and were accompanied by a necking phenomenon (Fig. 8). The transition from crazing to shear yielding was hypothesized to be due to poorly bonded nanoparticles that are well dispersed in the polymer matrix. In this way, the polymer was hypothesized to detach from the nanoparticles and deform in response to the tensile stress, thus delocalizing the yield damage and preventing the typical craze-to-fracture brittle failure.

The most dramatic increases in the modulus of nanocomposites occur in the region above the T_g. Often, these increases are much greater (4000%) than those that occur below T_g.[14,67] This is hypothesized to be due to the creation of crosslinks, either temporary or permanent, between nanoparticles and polymer which serve to increase the plateau modulus. Indeed, the ability to strictly control the size and surface activity of nanoparticles in recent dynamic mechanical studies on silica/polyvinylacetate nanocomposites above the T_g has shed new light on the nonlinear reinforcement behavior of rubbery melts.[68] In this study, the modulus of the polymer nanocomposite melt is shown to vary depending on the imposition of a large or small amplitude strain. The corresponding hypothesis concludes that the nanoparticles act as centers for trapped, nonlabile polymer chains that serve as temporary entanglements. In this work, the destruction and eventual recovery of the network offer an alternative to the particle filler networks widely believed to be responsible for the behavior known as the Payne effect.[69]

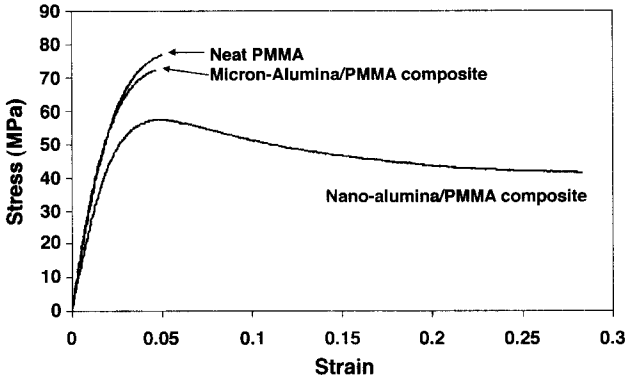


Fig. 7 Stress–strain curves for alumina/polymethyl methacrylate (PMMA) nanocomposites showing transformation from brittle to ductile behavior at 5 wt.% 39-nm-diameter alumina nanoparticles. (From Ref. [66].)

Studies of the wear behavior of nanofilled polymer composites have also demonstrated encouraging results. In composites using micron-size fillers, the wear resistance increases, but often the coefficient of friction does as well. In silica/epoxy nanocomposites at low volume fractions (~2 vol.%), however, sliding wear tests showed that both the frictional coefficient and the specific wear rate were lower than either the unfilled epoxy or epoxy filled with micron-size fillers.[70] This is significant because most filled systems display the decreased wear rate, but also show an increase in the friction coefficient. Nanoparticles are theorized to be able to significantly reduce wear rates because the wear mechanism changes from severe abrasion wear to mild sliding wear. This is due to the reduction in the angularity of the filler particles and the fact that the transfer film, where abrasion is taking place, can be strengthened because the nanoparticles would have the capability of blending well with wear particles. In addition, the material removal of nanoparticulate composites is thought to be much milder than that of conventional composites because the fillers have the same size as the segments of the surrounding polymer chains.[70]

The scratch resistance of nanocomposites is also improved over their micron counterparts.[71] Ng et al. reported that nano-TiO_2-filled epoxy showed greater scratch resistance with the scratch track being much smoother than either the neat polymer or a micron-filled system.[72] In addition, the depth and width of the scratch track decreased in the nanofilled system. This behavior has also been seen in alumina/gelatin nanocomposites.[73]

Nanoparticles have also been recently used to reinforce shape memory polymers for microelectromechanical systems (MEMS) applications. Gall et al.[74] demonstrated that the microhardness and elastic modulus could be increased through the use of SiC nanoparticles. In this case, the small size of the MEMS devices necessitates the use of nanoparticles for incorporation into the device, but the nanoparticles actually increased the force of recovery from 300 mN for the unfilled resin to 450 mN for a 20 wt.% nanofilled composite.

As with most mechanical systems, the property enhancements shown above experience an initial increase followed by a decrease usually in the 2–5 vol.% range. The observation of a particular weight fraction at which mechanical behavior is optimized has been observed in other nanocomposite studies.[14,62,65] It is obvious that at lower particle loadings, there is generally not enough material altered by the presence of the particles to permit large-scale changes in material properties. Higher loadings

induce agglomeration, which degrades the polymer performance through, for example, inclusion of voids that can act as preferential sites for crack initiation and failure.

Thermal Properties

The thermal properties of polymers and their composites are important from a mechanical stability perspective as well as a processing perspective. Thus the ability to alter this behavior, either by increasing or decreasing the processing or service temperature, is highly important in finding new markets for polymer products.

One of the most widely measured values to quantify thermal properties in polymers is the glass transition temperature. Changes in the glass transition temperature, T_g, as a function of the filler content have been reported for a variety of polymer composites containing a wide variety of nanoscale fillers. Most researchers report an increase in the T_g as a function of filler content;[10,14,75] in fact, Avella and coworkers have shown a 35°C increase in the T_g of PMMA filled with just 6 wt.% $CaCO_3$ nanoparticles.[76] On the other hand, decreases in the nanocomposite T_g have also been reported.[11,15]

The increases in T_g have been linked to the immobilization of the matrix within an IZ which hinders cooperative motion by raising the energy barrier for intermolecular chain movement.[11,75] However, attempts to relate the T_g reductions to existing theories of the glass transition have been limited. Becker et al.[11] reported that, over the range of filler volume fractions considered (0–10 vol.%), the T_g of their acetoxypropyltrimethoxysilane (APTS)-treated silica-filled PMMA–HEMA copolymer system decreased by 16°C, while the T_g of their composites prepared with as-received filler decreased by 9°C. This was contrasted with the increases in T_g observed for 10-nm silica coated with methacryloxypropyltrimethoxysilane (MPTS). In these latter specimens, the T_g increased by 10°C at 10 vol.% filler. However, the 100-nm silica used in the study produced no changes in the T_g with either coating at any filler concentration. Thus the increased surface area of the 10-nm silica and the interface between the APTS- and MPTS-coated particles and the polymer dramatically changed the polymer mobility in the bulk at fairly low volume fractions resulting in both increasing and decreasing T_g values.

In some instances, the particle, IZ, and matrix have been modeled as a core-shell scheme to incorporate the interaction zone/low mobility region. If the volume fraction of the interaction zone is high enough, this model shows two T_gs, while most researchers do not see this, there have been reports that show the additional relaxation mode corresponding to the restricted polymer network in contact with the nanofiller.[77,78]

Recent work by Sternstein and Zhu[68] on silica/polyvinylacetate nanocomposites indicates that a greater far-field effect is at work in altering chain dynamics than simply a local immobilization of the polymer chain as has been previously suggested.[78] This far-field effect may have a very large effect on the glass transition as shown recently by Ash et al.[15] In this work, the T_g of alumina/PMMA nanocomposites at 1.0 wt.% filler drops by 25°C when compared to the neat polymer (Fig. 8). Further additions of filler do not lead to additional T_g reductions. This novel thermal behavior is shown to vary with particle size, but this dependence can be normalized with respect to specific surface area. The nanocomposite T_g phenomenon is hypothesized to be due to nonadhering nanoparticles that act as well-dispersed internal void/polymer interfaces that break up the percolating structure of dynamically heterogeneous domains recently suggested by Long and Lequeux[79] to be responsible for the T_g reductions in polymer ultrathin films.[80] The results also point to a dramatically increased scope of the so-called IZ and its "far-field" effect on the entire matrix.

A closely related topic in polymers and nanocomposites is the diffusivity of these materials. Studying gas transport of nanosilica-filled polymer by 129Xe NMR permeativity increased with increasing filler content while retaining gas molecule size selectivity.[81] Similar results were shown using pulsed field gradient (PFG) NMR techniques with an order of magnitude increase in translational diffusion seen at 30 wt.% silica in a random copolymer.[82]

Optical Properties

As stated in the Introduction, the scattering of light by nanoscale particles is markedly reduced when the particle size is decreased less than 50 nm.[8] Thus nanoparticles may be added to polymers to increase wear, toughness, etc., while maintaining the optical transparency that many organic materials possess. An excellent review, both historical and current, of the phenomena of light absorption and scattering by nanoparticles in solution and in nanocomposites was recently given by Caseri.[8] Apart from maintaining clarity in the visible wavelengths, adding nanoparticles can add interesting functionality in the form of UV absorption[83] and the alteration of the refractive index of the matrix polymer. The addition of TiO_2 nanoparticles to many polymers results in such a material where the particles act to absorb the UV radiation and increase the refractive index.[71,84] Indeed, nanocomposites of lead sulfide and polyethylene oxide now boast a refractive index of 3.9,[3,85] while gold nanoparticles contained in gelatin have reduced the refractive index to about 1.[86] Dichoric behavior has also been observed in nanoparticle-filled polymers that have been highly

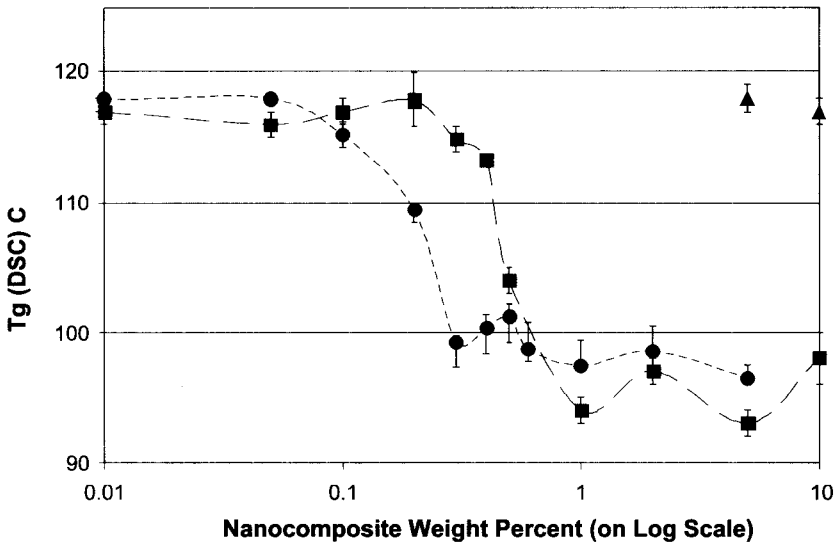


Fig. 8 Glass transition behavior of alumina/PMMA nanocomposites (■—38/39-nm alumina, ●—17-nm alumina). Note that the filler weight fraction is plotted on a log scale to show the behavior of the lower values more clearly. The neat PMMA is plotted as 0.01 wt.% on the above graph. Following coating with GPS, the T_g returns to the neat value (▲—GPS-coated). (From Ref. [15].)

stretched, resulting in "string of pearls"-type structures of fillers with very high aspect ratios. In these composites, the interaction with polarized light varies with the direction of both the incoming radiation and the orientation of the filler strands and results in both color[87] and intensity changes.[88] This type of composite response can be used in the liquid crystal displays[8] or in strain-based schemes for light-filtering devices.

Electrical Properties

The opportunities to tune the electrical properties of polymers with nanocomposites extend beyond enhancing conductive polymers; superparamagnetism in magnetic materials[89,90] or tunable band gaps in quantum dot semiconductor arrays show promise when incorporated into polymers for additional functionality. In addition, the percolation threshold has been shown to be lower in nanocomposites than in traditional filled composites with values as low as 0.03 vol.% reported.[91] Hong et al. using a ZnO/LDPE nanocomposite showed a lower percolation limit and a slower decrease in the resistivity with nanofiller concentration compared to conventional composite.[92] The dielectric breakdown strength of nanocomposites can also be enhanced over compositions with micron-scale fillers.[18]

In addition to traditional filler roles, organic–inorganic nanocomposites have found their way into etch resists used in semiconductor nanolithography. With line sizes in CMOS technology approaching 100 nm, adding silica nanoparticles to traditional organic resists, such as PMMA, have been shown to increase rigidity and display a higher glass transition temperature which enhances resist performance for nanometer pattern fabrication.[93]

CARBON NANOTUBE/POLYMER NANOCOMPOSITES

Carbon nanotubes are in some ways the ideal fiber. The almost defect-free structure of the nanotubes results in mechanical properties that are comparable to those of a graphene sheet. Recent studies have focused on the assessment of the mechanical properties of carbon nanotubes through both experiment and modeling. The Young's modulus of these almost-perfect fibers has been predicted to be between 0.64 and 5.0 TPa for a SWNT.[94–100] The tensile strength of SWNTs was calculated to range from 50 to 100 GPa depending on the type and number of defects present.[101] The nanometer-scale size of the nanotubes introduces technical difficulties in the experimental measurement of their mechanical properties. However, Yu and coworkers conducted tensile testing experiments on individual MWNTs that were attached to opposing AFM tips on both ends.[102] They obtained values of 11–63 GPa for the tensile strength and 270–950 GPa for the elastic modulus of the MWNT. Yu and coworkers also conducted tensile tests on SWNT bundles and obtained 320–1470 GPa for the elastic modulus.[103] Thus the strengths obtained by experimental measurements are lower than the calculated values, possibly due to intrinsic nanotube defects that cause premature

mechanical failure or to other errors in the experimental methodology. Indirect measurements have yielded similar values for modulus and strength.[104,105]

In addition to their unique mechanical properties, carbon nanotubes have a high aspect ratio, high thermal conductivity, high electrical conductivity, and low coefficient of thermal expansion.[106] In short, carbon nanotubes theoretically have the properties of an ideal graphite fiber. This alone makes them attractive for incorporation into polymers. In addition, their small size introduces the possibility of synthesizing transparent materials with low volume fractions of filler.

Mechanical Properties

Studies have been conducted in order to explore the mechanical properties of nanotube-reinforced polymer composites, keeping in mind the challenges to be addressed. MWNT–polystyrene films were prepared from solution in toluene and tested in tension.[28] 1 wt.% of MWNT caused an increase of 36% in the elastic modulus, and 25% increase in strength. In another study, 5 wt.% MWNT–epoxy samples were prepared and tested under tension and compression.[107] An increase of 20% in the tensile elastic modulus and the compression elastic modulus was observed. Insights into the mechanisms of load transfer were discussed for this system, because of the interesting behavior of the composites under tension and compression, as measured by Raman spectroscopy (see next section for details). In another study, carbon nanofiber reinforced poly(etherether)ketone (PEEK) nanocomposites were prepared by twin-screw extrusion and tested by dynamic mechanical analysis and tensile tests.[108] A gradual increase in the elastic modulus as a function of nanofiber concentration was observed such that a loading of 15 wt.% nanofibers resulted in a 40% increase in the modulus. The effect of noncovalent modification of carbon nanotubes on the mechanical and thermal properties of nanotube-epoxy composite was investigated[37] using polyoxyethylene 8 lauryl as a surfactant. In this work by Gong et al. an increase of 25°C in the glass transition temperature and a 30% increase in the elastic modulus, as measured by dynamic mechanical analysis, were observed with the incorporation of only 1 wt.% nanotubes. Using polyvinylidene fluoride, again as a noncovalent modifier, in MWNT–polymethyl methacrylate (PMMA) nanocomposites[109] the dispersion of the nanotubes in the polymer was better, and the storage modulus, as measured by dynamic mechanical analysis, increased. In situ polymerization has been attempted also in the presence of carbon nanotubes, in order to achieve the participation of the nanotubes in the polymerization process and hence get a covalent attachment of the growing polymer chains to the nanotube surface. Such a study was conducted with methyl methacrylate polymerization and MWNT.[110] Evidence for nanotube–polymer interactions was observed from the examination of the fracture surface of the composites. An increase in the heat deflection temperature was also observed with the incorporation of nanotubes. This in situ polymerization procedure was also used in a SWNT–polyimide system.[111] In this research the electrical conductivity of the nanocomposite increased by 10 orders of magnitude at 0.2 wt.% SWNT loading while the mechanical properties also improved. Fluorinated MWNT-reinforced polyvinyl alcohol composites were prepared and mechanically tested.[112] A 400% increase in the storage modulus (as measured by dynamic mechanical analysis) was obtained for 4 wt.% functionalized MWNT composite as compared to pure polyvinyl alcohol.

From a mechanical properties viewpoint, the characteristic bending and looping of SWNTs is problematic if the maximum reinforcement is to be obtained. Therefore research has been conducted to align the nanotubes by magnetic fields[113] and by melt spinning.[114] In both cases nanocomposites with anisotropic properties were obtained with increases in the elastic modulus realized in the alignment orientation.

Few if any of the nanotube-filled polymer composites reported in the literature have achieved the properties expected by their high modulus and strength. This is because of three closely related factors: dispersion, load transfer, and volume fraction. For example, SWNTs are synthesized in bundles. In the bundle form, the load-bearing portion of the bundle is only the outermost tubes. The tubes inside the bundle slide are relative to the others and do not carry significant load. To solve this problem, the SWNT must be separated from the bundles and as mentioned earlier progress is being made in this area. MWNTs on the other hand, can be separated from their bundles. MWNTs however, are concentric tubes with poor bonding between the carbon layers. Therefore even if the outermost tube of a MWNT is carrying load, the inner layers may not be. This also reduces their efficiency. Finally, the high surface area of the nanotubes makes it difficult to process composites with high volume fractions of fillers because the viscosity of the nanotube/matrix mixture is prohibitively high.

Electrical and Optical Properties

One likely application of nanotube-filled polymers is for applications requiring high electrical conductivity. Their high aspect ratio leads to percolation at extremely low loadings. A recent paper claimed percolation at 0.0025 wt.%[115] and several others report levels near 0.3 wt.%.[116] This low percolation loading and relatively high conductivity leads to the promise of transparent conducting coatings or other multifunctional applications where the conductivity is improved, but other properties are not compromised. Changes in conductivity as high as 8 orders

of magnitude have been observed in nanotube/PPV composites with little loss in photoluminescence/electroluminescence yield. In addition, the mechanical strength simultaneously improved.[117] In situ polymerization under sonication has also been used to assist the dispersion of SWNT in a polyimide matrix. Enhancement of the conductivity by 10 orders of magnitude was measured in this system for 0.2 wt.% of SWNT, while maintaining transparency of the nanocomposite.[111] Alignment has also been found to improve conductivity.[114] The conductivity that can be achieved is somewhat limited by the mixture of semiconducting and metallic nanotubes produced. Improvements in the ability to separate metallic from semiconducting nanotubes will likely lead to further improvements in conductivity. Other electrical properties such as permittivity have also been investigated.[118]

The conduction properties of nanotubes can also be used to alter the optical response of polymers. For example, functionalization of nanotubes and doping of chemically modified nanotubes in low concentration to photoactive polymers such as PPV have been shown as a means to alter the hole transport mechanism and hence the optical emission of the polymer.[119] Such hole-collecting properties seem to be a quite general property of some conjugated polymers. Specifically, SWNTs in a polymer matrix trap holes injected from the anode in organic light emitting diodes (OLEDs).[120]

The optical limiting behavior of polymer nanotube-filled composites has been investigated with success. It has been reported that polymerization of phenylacetylene in the presence of short nanotubes results in a solvated product of tubules wrapped with poly(phenylacetylene). The optical limiting properties of this composite are excellent and far better than the pristine polymer and polymer filled with other carbon structures. They also fare well in photovoltaic applications suggesting that such nanocomposites could find an array of potential applications in optics-related and laser-based technologies.[121]

NANOTUBE AND NANOPARTICLE COMPOSITE MODELS

The dramatic changes in mechanical properties observed in nanofillers challenge existing theories of mechanical reinforcement. In traditional composites, the modulus can be predicted, usually between upper and lower bounds,[122] by modeling the effectiveness of load transfer from the matrix to the polymer. As the particles are, in large part, stiffer than the surrounding matrix, this load transfer results in a stiffer composite than the neat polymer. With the nanoparticles, however, the "span" required for effective load transfer is extremely small and thus can contribute only a small portion of the dramatic gains in modulus seen in nanocomposites. In addition, because of their unique geometry, the theoretical modeling of the mechanical behavior of nanotube polymer composites is even more complex. The hollow shape of the nanotubes, the wall–wall interactions in MWNT, and the molecular level interactions between the nanotubes and the polymer matrix have to be considered in these models. Odegard and coworkers[123] transformed the SWNT and the neighboring polymer chains into an effective fiber. By using molecular dynamic considerations, the effects of nanotube alignment and length on the elastic modulus were calculated. These theoretical calculations were compared to the experimental results of SWNT–polyimide composite mechanical properties.

Addressing the presence of a large volume fraction of the IZ that has properties different from the bulk polymer is also paramount to successful modeling. The alteration of the matrix that takes place at the IZ, as previously mentioned, is a likely candidate for property enhancement. Modeling the IZ is difficult; however, as imperfect bonding, mechanical stresses, and chemical interactions can all play a decisive role in determining its behavior. Ji et al.[124] recently modified a composite model to incorporate the interphase zone and were able to successfully model the increases in the modulus of a clay/nylon-6 nanocomposite system. Their treatment modeled the interphase zone as a linear gradient change in the modulus between the surface of the filler and the matrix and introduced the critical effect of particle size (through the increase in volume fraction of the interphase) into the theory. The curves predicted by the model were shown to be in good agreement with the experimental nanocomposite mechanical behavior.

CONCLUSION

Polymer nanocomposites incorporating carbon nanotubes or nanoparticles are a novel class of composite materials that are often multifunctional, adding the unique optical, electrical, or mechanical properties of the nanofillers while maintaining, if not enhancing, the neat polymer properties. These composites are often characterized by low volume fractions of filler, which may then affect the entire matrix, due to their large surface-area-to-volume ratio, through an interaction zone. The unique nature of these particles presents both challenges and unique opportunities to create a wide range of multifunctional polymer composites.

REFERENCES

1. Bueche, A.M. Filler reinforcement of silicone rubber. J. Polym. Sci. **1957**, *25*, 139–149.

2. Kuriakose, B.; De, S.K.; Bhagawan, S.S.; Sivaramkrishnan, R.; Athithan, S.K. Dynamic mechanical properties of thermoplastic elastomers from polypropylene–natural rubber blend. J. Appl. Polym. Sci. **1986**, *32*, 5509–5521.
3. Sumita, M.; Shizuma, T.; Miyasaka, K.; Shikawa, K.J. Effect of reducible properties of temperature, rate of strain, and filler content on the tensile yield stress of nylon 6 composites filled with ultrafine particles. Macromol. Sci.-Phys. **1983**, *B22*, 599–616.
4. Sumita, M.; Tsukumo, Y.; Miyasaka, K.; Ishikawa, K. Tensile yield stress of polypropylene composites filled with ultrafine particles. J. Mater. Sci. **1983**, *18*, 1758–1764.
5. Li, T.; Chen, Q.; Schadler, L.S.; Siegel, R.W.; Mendel, J.; Irvin, G.C., Jr. Scratch behavior of nanophase Al_2O_3-filled gelatin films. Polym. Compos. **2002**, *23* (6), 1076–1086.
6. Ruan, S.L.; Gao, P.; Yang, X.G.; Yu, T.X. Toughening high performance ultrahigh molecular weight polyethylene using multiwalled carbon nanotubes. Polymer **2003**, *44*, 5643–5654.
7. Ash, B.J.; Stone, J.; Rogers, D.F.; Schadler, L.S.; Siegel, R.W.; Benicewicz, B.C.; Apple, T. Investigations into the thermal and mechanical behavior of PMMA/alumina nanocomposites. Mater. Res. Soc. Symp. Proc. **2001**, *661*, KK2.10.1–KK2.10.6.
8. Caseri, W. Nanocomposites of polymers and metals or semiconductors: Historical background and optical properties. Macromol. Rapid Commun. **2000**, *21*, 705–722.
9. Siegel, R.W. Nanophase Materials. In *Encyclopedia of Applied Physics*; Trigg, G.L., Ed.; Weinheim: VCH Publishing, 1994; Vol. 11.
10. Hergeth, W.; Steinau, U.; Bittrich, H.; Simon, G.; Schmutzler, K. Polymerization in the presence of seeds: Part IV. Emulsion polymers containing inorganic filler particles. Polymer **1989**, *30*, 254–258.
11. Becker, C.; Mueller, P.; Schmidt, H. Optical and thermomechanical investigations on thermoplastic nanocomposites with surface modified silica nanoparticles. SPIE **1998**, *3469*, 88–98.
12. Carotenuto, G.; Nicolais, L.; Kuang, X.; Zhu, Z. A method for the preparation of PMMA–SiO_2 nanocomposites with high homogeneity. Appl. Compos. Mater. **1995**, *2*, 385–393.
13. Wang, Z.; Pinnavaia, T.J. Hybrid organic–inorganic nanocomposites: Exfoliation of magadiite nanolayers in an elastomeric epoxy polymer. Chem. Mater. **1998**, *10*, 1820–1826.
14. Becker, C.; Krug, H.; Schmidt, H. Tailoring of thermomechanical properties of thermoplastic nanocomposites by surface modification of nanoscale silica particles. Mater. Res. Soc. Symp. Proc. **1996**, *435*, 237–242.
15. Ash, B.J.; Schadler, L.S.; Siegel, R.W. Glass transition behavior of alumina/polymethylmethacrylate nanocomposites. Mater. Lett. **2002**, *55*, 83–87.
16. Ding, W.; Eitan, A.; Fisher, F.T.; Chen, X.; Dikin, D.A.; Andrews, R.; Brinson, L.C.; Ruoff, R.S. Direct observation of polymer sheathing in carbon nanotube—Polycarbonate composites. Nano Lett. **2003**, *3* (11), 1593–1587.
17. Chan, C.M.; Wu, J.; Li, J.X.; Cheung, Y.K. Polypropylene/calcium carbonate nanocomposites. Polymer **2002**, *43*, 2981–2992.
18. Ma, D.L.; Onnesby, C.; Martensson, E.; Siegel, R.W.; Schadler, L.S. Influence of nanoparticle surface water on the breakdown strength of nanoparticle filled low density polyethylene. **2003**.
19. Bhimaraj, P.; Freudenberg, K.; Action, J.; Sawyer, G.; Schadler, L. Enhancement in wear resistance and self-lubrication of PET through addition of nanofillers. **2003**.
20. Abboud, M.; Turner, M.; Duguet, E.; Fontanille, M. PMMA-based composite materials with reactive ceramic fillers: Part 1. Chemical modification and characterization of ceramic particles. J. Mater. Chem. **1997**, *7*, 1527.
21. Wang, S.; Wang, M.; Lei, Y.; Zhang, L. Anchor effect of poly(styrene maleic anhydride)/TiO_2 nanocomposites. J. Mater. Sci. Lett. **1999**, *18*, 2009–2012.
22. Caris, C.; van Elven, L.; van Herk, A.; German, A. Polymerization of MMA (methyl methacrylate) at the surface of inorganic submicron particles. Br. Polym. J. **1989**, *21*, 133–140.
23. Rong, M.Z.; Zhang, M.Q.; Zheng, Y.X.; Zeng, H.M.; Friedrich, K. Improvement of tensile properties of nano-SiO_2/PP composites in relation to percolation mechanism. Polymer **2001**, *42*, 3301–3304.
24. von Werne, T.; Patten, T. Atom transfer radical polymerization from nanoparticles: A tool for the preparation of well-defined hybrid nanostructures and for understanding the chemistry of controlled/'living' radical polymerizations from surfaces. J. Am. Chem. Soc. **2001**, *123*, 7497–7505.
25. Weimer, M.W.; Chen, H.; Giannelis, E.P.; Sogah, D.Y. Direct synthesis of dispersed nanocomposites by in situ living free radical polymerization using a silicate-anchored initiator. J. Am. Chem. Soc. **1999**, *121*, 1615–1616.

26. Tsubokawa, N.; Kogure, A.; Maruyama, K.; Sone, Y.; Shimomura, M. Graft polymerization of vinyl monomers from inorganic ultrafine particles initiated by azo groups introduced onto the surface. Polym. J. **1990**, *22*, 827–833.
27. Caruso, F. Nanoengineering of particle surfaces. Adv. Mater. **2001**, *13*, 11–22.
28. Qian, D.; Dickey, E.C.; Andrews, R.; Rantell, T. Load transfer and deformation mechanisms in carbon nanotube–polystyrene composites. Appl. Phys. Lett. **2000**, *76* (20), 2868–2870.
29. Dalton, A.B.; Bayrne, H.J.; Coleman, J.N.; Curran, S.; Davey, A.P.; McCarthy, B.; Blau, W. Optical absorption and fluorescence of multi-walled nanotube–polymer composites. Synth. Met. **1999**, *102*, 1176–1177.
30. Kahn, M.G.C.; Banerjee, S.; Wong, S.S. Solubilization of oxidized single-walled carbon nanotubes in organic and aqueous solvents through organic derivatization. Nano Lett. **2002**, *2* (11), 1215–1218.
31. Bower, C.; Kleinhammes, A.; Wu, Y.; Zhou, O. Intercalation and partial exfoliation of single-walled carbon nanotubes by nitric acid. Chem. Phys. Lett. **1998**, *288*, 481–486.
32. Nagasawa, S.; Yudasaka, M.; Hirahara, K.; Ichihashi, T.; Ijima, S. Effect of oxidation on single-wall carbon nanotubes. Chem. Phys. Lett. **2000**, *328*, 374–380.
33. Chen, J.; Hamon, M.A.; Hu, H.; Chen, Y.; Rao, A.M.; Eklund, P.C.; Haddon, R.C. Solution properties of single-walled carbon nanotubes. Science **1998**, *282*, 95–98.
34. Niyogi, S.; Hamon, M.A.; Hu, H.; Zhao, B.; Bhowmik, B.; Sen, R.; Itkis, M.E.; Haddon, R.C. Chemistry of single-walled carbon nanotubes. Acc. Chem. Res. **2002**, *35*, 1105–1113.
35. Carrillo, A.; Swartz, J.A.; Gamba, J.M.; Kane, R.S.; Chakrapani, N.; Wei, B.; Ajayan, P. Noncovalent functionalization of graphite and carbon nanotubes with polymer multilayers and gold nanoparticles. Nano Lett. **2003**, *3* (11), 1437–1440.
36. Chen, J.; Liu, H.; Weimer, W.A.; Halls, M.D.; Waldeck, D.H.; Walker, G.C. Noncovalent engineering of carbon nanotube surfaces by rigid, functional conjugated polymers. J. Am. Chem. Soc. **2002**, *124* (31), 9034–9035.
37. Gong, X.; Liu, J.; Baskaran, S.; Voise, R.D.; Young, J.S. Surfactant-assisted processing of carbon nanotube/polymer composites. Chem. Mater. **2000**, *12*, 1049–1052.
38. O'Connell, M.J.; Boul, P.; Ericson, L.M.; Huffman, C.; Wang, Y.; Haroz, E.; Kuper, C.; Tour, J.; Ausman, K.D.; Smalley, R.E. Reversible water-solubilization of single-walled carbon nanotubes by polymer wrapping. Chem. Phys. Lett. **2001**, *342*, 265–271.
39. Czerw, R.; Guo, Z.; Ajayan, P.M.; Sun, T.-P.; Carroll, D.L. Organization of polymers onto carbon nanotubes: A route to nanoscale assembly. Nano Lett. **2001**, *1* (8), 423–427.
40. McCarthy, B.; Coleman, J.N.; Crezw, R.; Dalton, A.B.; in het Panhuis, M.; Drury, A.; Brenier, P.; Nagy, J.B.; Lahr, B.; Byrne, H.J.; Carroll, D.L.; Blau, W.J. A microscopic and spectroscopic study of interactions between carbon nanotubes and a conjugated polymer. J. Phys. Chem., B **2002**, *106* (9), 2210–2216.
41. Hamon, M.A.; Hu, H.; Bhowmik, P.; Niyogi, S.; Zhao, B.; Itkis, M.E.; Haddon, R.C. End-group and defect analysis of soluble single-walled carbon nanotubes. Chem. Phys. Lett. **2001**, *347*, 8–12.
42. Lin, Y.; Rao, A.M.; Sadanadan, B.; Kenik, E.A.; Sun, Y.-P. Functionalization of multiple-walled carbon nanotubes with aminopolymers. J. Phys. Chem., B **2002**, *106* (6), 1294–1298.
43. Hill, D.E.; Rao, A.M.; Allard, L.F.; Sun, Y.-P. Functionalization of carbon nanotubes with polystyrene. Macromolecules **2002**, *35* (25), 9466–9471.
44. Zhao, W.; Song, C.; Pehrsson, P.E. Water-soluble and optically PH-sensitive single-walled carbon nanotubes from surface modification. J. Am. Chem. Soc. **2002**, *124* (42), 12418–12419.
45. Liu, J.; Rinzler, A.G.; Dai, J.H.; Hafner, J.H.; Bradley, R.K; Boul, P.J.; Lu, A.; Iverson, T.; Shelimov, K.; Huffman, C.B.; Rodrigez-Macias, F.; Shon, Y.S.; Lee, T.R.; Colbert, D.T.; Smalley, R.E. Fullerene pipes. Science **1998**, *280*, 1253–1256.
46. Hamon, M.A.; Chen, J.; Hu, H.; Chen, Y.; Itkis, M.E.; Rao, A.M.; Eklund, P.C.; Haddon, R.C. Dissolution of single-walled carbon nanotubes. Adv. Mater. **1999**, *11* (10), 834–840.
47. Baker, S.E.; Cai, W.; Lasseter, T.L.; Weidkamp, K.P.; Hamers, R.J. Covalently bonded adducts of deoxyribonucleic acid (DNA) oligonucleotides with single-wall carbon nanotubes: Synthesis and hybridization. Nano Lett. **2002**, *2* (12), 1413–1417.
48. Huang, W.; Fernando, S.; Allard, L.F.; Sun, Y.-P. Solubilization of single-walled carbon nanotubes with diamine-terminated oligomeric poly(ethylene glycol) in different functionalization reactions. Nano Lett. **2003**, *3* (4), 565–568.
49. Huang, W.; Taylor, S.; Fu, K.; Lin, Y.; Zhang, D.; Hanks, T.W.; Rao, A.M.; Sun, Y.-P. Attaching

proteins to carbon nanotubes via diimide-activated amidation. Nano Lett. **2002**, *2* (4), 311–314.
50. Banerjee, S.; Wong, S.S. Synthesis and characterization of carbon nanotube–nanocrystal heterostructures. Nano Lett. **2002**, *2* (3), 195–200.
51. Mickelson, E.T.; Huffman, C.B.; Rinzler, A.G.; Smalley, R.E.; Hauge, R.H.; Margrave, J.L. Fluorination of single-wall carbon nanotubes. Chem. Phys. Lett. **1998**, *296*, 188–194.
52. Ying, Y.; Saini, R.K.; Liang, F.; Sadana, A.K.; Billups, W.E. Functionalization of carbon nanotubes by free radicals. Org. Lett. **2003**, *5* (9), 1471–1473.
53. Holzinger, M.; Votrowsky, O.; Hirsch, A.; Hennrich, F.; Kappes, M.; Weiss, R.; Jellen, F. Sidewall functionalization of carbon nanotubes. Angew. Chem., Int. Ed. **2001**, *40* (21), 4002–4005.
54. Banerjee, S.; Wong, S.S. Structural characterization, optical properties, and improved solubility of carbon nanotubes functionalized with Wilkinson's catalyst. J. Am. Chem. Soc. **2002**, *124*, 8940–8948.
55. Georgakilas, V.; Kordatos, K.; Prato, M.; Guldi, D.M.; Holzinger, M.; Hirsch, A. Organic functionalization of carbon nanotubes. J. Am. Chem. Soc. **2002**, *124* (5), 760–761.
56. Dyke, C.A.; Tour, J.M. Solvent-free functionalization of carbon nanotubes. J. Am. Chem. Soc. **2003**, *125*, 1156–1157.
57. Velasco-Santos, C.; Martínez-Hernández, A.L.; Lozada-Cassou, M.; Alvarez-Castillo, A.; Castaño, V.M. Chemical functionalization of carbon nanotubes through an organosilane. Nanotechnology **2002**, *13*, 495–498.
58. Bahr, L.J.; Tour, J.M. Highly functionalized carbon nanotubes using in situ generalized diazonium compounds. Chem. Mater. **2001**, *13* (11), 3823–3824.
59. Ajayan, P.M.; Braun, P.; Schadler, L.S. *Nanocomposite Science and Technology*; Wiley-VCH Verlag GmbH & Co. KgaA: Weinheim, Germany, 2003.
60. Hussain, M.; Nakahara, A.; Nishijima, S.; Niihara, K. Fracture behavior and fracture toughness of particulate filled epoxy composites. Mater. Lett. **1996**, *27*, 21–25.
61. Sumita, M.; Tsukihi, H.; Miyasaka, K.; Ishikawa, K. Dynamic mechanical properties of polypropylene composites filled with ultrafine particles. J. Appl. Polym. Sci. **1984**, *29*, 1523–1530.
62. Ou, Y.; Yang, F.; Yu, Z. A new conception on the toughness of nylon 6/silica nanocomposite prepared via in situ polymerization. J. Polym. Sci., B, Polym. Phys. **1998**, *36*, 789–795.
63. Petrovic, Z.S.; Zhang, W. Glassy and elastomeric polyurethanes filled with nano-silica particles. Mat. Sci. Forum **2000**, *352*, 171–176.
64. Petrovic, Z.S.; Javni, I.; Waddon. A.; Banhegyi, G. Structure and properties of polyurethane–silica nanocomposites. J. Appl. Polym. Sci. **2000**, *76*, 133–151.
65. Zhang, J.; Wang, X.; Lu, L.; Li, D.; Yang, X. Preparation and performance of high-impact polystyrene (HIPS)/nano-TiO_2 nanocomposites. J. Appl. Polym. Sci. **2003**, *87*, 381–385.
66. Ash, B.J.; Rogers, D.F.; Wiegand, C.J.; Schadler, L.S.; Siegel, R.W.; Benicewicz. B.C.; Apple, T. Mechanical properties of Al_2O_3/polymethylmethacrylate nanocomposites. Polym. Compos. **2002**, *23*, 1014–1025.
67. Zhang, Q.; Archer, L.A. Poly(ethylene oxide)/silica nanocomposites: Structure and rheology. Langmuir **2002**, *18*, 10435–10442.
68. Sternstein, S.S.; Zhu, A.-J. Reinforcement mechanism of nanofilled polymer melts as elucidated by nonlinear viscoelastic behavior. Macromolecules **2002**, *35*, 7262–7273.
69. Payne, A.R. The dynamic properties of carbon black-loaded natural rubber vulcanizates. J. Appl. Polym. Sci. **1962**, *6*, 368–372.
70. Zhang, M.Q.; Rong, M.Z.; Yu, S.L.; Wetzel, B.; Friedrich, K. Effect of particle surface treatment on the tribological performance of epoxy based nanocomposites. Wear **2002**, *253*, 1086–1093.
71. Hwang, D.K.; Moon, J.H.; Shul, Y.G.; Jung, K.T.; Kim, D.H.; Lee, D.W. Scratch resistant and transparent UV-protective coating on polycarbonate. J. Sol-Gel Sci. Technol. **2003**, *26*, 783–787.
72. Ng, C.B.; Schadler, L.S.; Siegel, R.W. Synthesis and mechanical properties of TiO_2–epoxy nanocomposites. J. Nanostruct. Mater. **1999**, *12* (1–4), 507–510.
73. Li, T.; Chen, Q.; Schadler, L.S.; Siegel, R.W. Scratch behavior of nanoparticle Al_2O_3-filled gelatin films. Polym. Compos. **2002**, *23*, 1076–1086.
74. Gall, K.; Dunn, M.L.; Liu, Y.; Finch, D.; Lake, M.; Munshi, N.A. Shape memory polymer nanocomposites. Acta Mater. **2002**, *50*, 5115–5126.
75. Iisaka, K.; Shibayama, K. Mechanical alpha-dispersion and interaction in filled polystyrene and polymethylmethacrylate. J. Appl. Polym. Sci. **1978**, *22*, 3135–3143.
76. Avella, M.; Errico, M.E.; Martuscelli, E. Novel PMMA/$CaCO_3$ nanocomposites abrasion resistant prepared by an in situ polymerization process. Nano Lett. **2001**, *1*, 213–217.
77. Arrighi, V.; McEwen, I.J.; Qian, H.; Serrano Prieto, M.B. The glass transition and interfacial layer in

styrene–butadiene rubber containing silica nanofiller. Polymer **2003**, *44*, 6259–6266.
78. Tsagaropoulos, G.; Eisenberg, A. Direct observation of two glass transitions in silica-filled polymers. Implications for the morphology of random ionomers. Macromolecules **1995**, *28*, 396–398.
79. Long, D.; Lequeux, F. Heterogeneous dynamics at the glass transition in van der Waals liquids, in the bulk and in thin films. Eur. Phys. J., E **2001**, *4*, 371.
80. Mattsson, J.; Forrest, J.A.; Borgesson, L. Quantifying glass transition behavior in ultrathin freestanding polymer films. Phys. Rev., E. **2000**, *62*, 5187–5200.
81. Merkel, T.C.; Toy, L.G.; Andrady, A.L.; Gracz, H.; Stejskal, E.O. Investigation of enhanced free volume in nanosilica-filled poly(1-trimethylsilyl-1-propyne) by ^{129}Xe NMR spectroscopy. Macromolecules **2003**, *36*, 353–358.
82. Zhong, J.; Wen, W.Y.; Jones, A.A. Enhancement of diffusion in a high-permeability polymer by the addition of nanoparticles. Macromolecules **2003**, *36*, 6430–6432.
83. Kyprianidou-Leodidou, T.; Margraf, P.; Caseri, W.; Suter, U.W.; Walther, P. Polymer sheets with a thin nanocomposite layer acting as a UV filter. Polym. Adv. Technol. **1997**, *8*, 505–512.
84. Nussbaumer, R.J.; Caseri, W.R.; Smith, P.; Tervoort, T. Polymer–TiO_2 nanocomposites: A route towards visually transparent broadband UV filters and high refractive index materials. Macromol. Mater. Eng. **2003**, *288*, 44–49.
85. Weibel, M.; Caseri, W.; Suter, U.W.; Kiess, H.; Wehrli, E. Polymer nanocomposites with 'ultrahigh' refractive index. Polym. Adv. Technol. **1991**, *2*, 75–80.
86. Zimmermann, L.; Weibel, M.; Caseri, W.; Suter, U.W.; Walther, P. Polymer nanocomposites with 'ultralow' refractive index. Polym. Adv. Technol. **1993**, *4*, 1–7.
87. Dirix, Y.; Darribere, C.; Heffels, W.; Bastiaansen, C.; Caseri, W.; Smith, P. Optically anisotropic polyethylene–gold nanocomposites. Appl. Opt. **1999**, *38*, 6581–6586.
88. Ash, B.J. Investigation into the Thermal and Mechanical Behavior of Alumina/Polymethylmethacrylate Nanocomposites. Ph.D. Thesis; Rensselaer Polytechnic Institute, 2003; 108–113.
89. Roy, S.; Das, D.; Chakravorty, D.; Agrawal, D.C. Magnetic properties of glass–metal nanocomposites prepared by the sol–gel route and hot pressing. J. Appl. Phys. **1993**, *74*, 4746–4749.
90. Morup, S.; Tronc, E. Superparamagnetic relaxation of weakly interacting particles. Phys. Rev. Lett. **1994**, *72*, 3278–3281.
91. Banerjee, P.; Mandel, B.M. Conducting polyaniline nanoparticle blends with extremely low percolation thresholds. Macromolecules **1995**, *28*, 3940–3943.
92. Hong, J.I.; Schadler, L.S.; Siegel, R.W.; Martensson, E. Rescaled electrical properties of ZnO/LDPE nanocomposites. Appl. Phys. Lett. **2003**, *82*, 1956–1958.
93. Gonsalves, K.E.; Merhari, L.; Wu, H.; Hu, Y. Organic–inorganic nanocomposites: Unique resists for nanolithography. Adv. Mater. **2001**, *13*, 703–714.
94. Treacy, M.M.J.; Ebbesen, T.W.; Gibson, J.M. Exceptionally high young's modulus for individual carbon nanotubes. Nature **1996**, *381* (6584), 678–680.
95. Gao, G.H.; Cagin, T.; Goddard, W.A. Energetics, structure, mechanical and vibrational properties of single-walled carbon nanotubes. Nanotechnology **1998**, *9* (3), 184–191.
96. Lu, J.P. Elastic properties of carbon nanotubes and nanoropes. Phys. Rev. Lett. **1997**, *79* (7), 1297–1300.
97. Sanchez-Portal, D.; Artacho, E.; Soler, J.M.; Rubio, A.; Ordejon, P. Ab initio structural, elastic, and vibrational properties of carbon nanotubes. Phys. Rev., B **1999**, *59*, 12677–12688.
98. Yakobson, B.I.; Barber, C.J.; Bemhole, J. Nanomechanics of carbon tubes: Instabilities beyond linear response. Phys. Rev. Lett. **1996**, *76* (14), 2511–2514.
99. Zhou, X.; Zhou, J.J.; Ou-Yang, Z.C. Strain energy and Young's modulus of single-wall carbon nanotubes calculated from electronic energy-band theory. Phys. Rev., B **2000**, *62* (20), 13692–13696.
100. Popov, V.N.; Van Doren, V.E.; Balkanski, M. Elastic properties of single-walled carbon nanotubes. Phys. Rev., B **2000**, *61* (4), 3078–3084.
101. Belytschko, T.; Xiao, S.P.; Schatz, G.C.; Ruoff, R.S. Atomistic simulations of nanotube fracture. Phys. Rev., B **2002**, *65* (23), 235430/1–8.
102. Yu, M.F.; Lourie, O.; Dyer, M.J.; Moloni, K.; Kelly, T.F.; Ruoff, R.S. Strength and breaking mechanism of multiwalled carbon nanotubes under tensile load. Science **2000**, *287*, 637–640.
103. Yu, M.F.; Files, B.S.; Arepalli, S.; Ruoff, R.S. Tensile loading of ropes of single wall carbon nanotubes and their mechanical properties. Phys. Rev. Lett. **2000**, *84* (24), 5552–5555.
104. Krishnan, A.; Dujardin, E.; Ebessen, T.W.; Yanilos, P.N.; Treacy, M.M.J. Young's modulus of single-walled nanotubes. Phys. Rev., B **1998**, *58*, 14013.

105. Lourie, O.; Wagner, H.D. Evaluation of Young's modulus of carbon nanotubes by micro-Raman spectroscopy. J. Mater. Res. **1998**, *13* (9), 2418–2422.
106. Thostenson, E.T.; Ren, Z.F.; Chou, T.W. Advances in the science and technology of carbon nanotubes and their composites: A review. Compos. Sci. Technol. **2001**, *61* (13), 1899–1912.
107. Schadler, L.S.; Giennaris, S.C.; Ajayan, P.M. Load transfer in carbon nanotube epoxy composites. Appl. Phys. Lett. **1998**, *73* (26), 3842–3844.
108. Sandler, J.; Werner, P.; Shaffer, M.S.P.; Demchuk, V.; Altstädt, V.; Windle, A.H. Carbon-nanofibre-reinforced poly(ether ether ketone) composites. Compos., Part A **2002**, *33*, 1033–1039.
109. Jin, Z.; Pramoda, K.P.; Goh, S.H.; Xu, G. Poly(vinylidene fluoride)-assisted melt-blending of multi-walled carbon nanotube/poly(methyl methacrylate) composites. Mater. Res. Bull. **2002**, *37*, 271–278.
110. Jia, Z.; Wang, Z.; Xu, C.; Liang, J.; Wei, B.; Wu, D.; Zhu, S. Study on poly(methyl methacrylate) carbon nanotube composites. Mater. Sci. Eng., A **1999**, *271*, 395–400.
111. Park, C.; Ounaies, Z.; Watson, K.A.; Crooks, R.E.; Smith, J., Jr.; Lowther, S.E.; Connell, J.W.; Siochi, E.J.; Harrison, J.S.; St. Clair, T.L. Dispersion of single wall carbon nanotubes by in situ polymerization under sonication. Chem. Phys. Lett. **2002**, *364*, 303–308.
112. Geng, H.Z.; Rosen, R.; Zheng, B.; Shimoda, H.; Fleming, L.; Liu, J.; Zhou, O. Fabrication and properties of poly(ethylene oxide) and functionalized carbon nanotubes. Adv. Mater. **2002**, *14* (19), 1387–1390.
113. Kimura, T.; Ago, H.; Tobita, M.; Ohshima, S.; Kyotani, M.; Yumura, M. Polymer composites of carbon nanotubes aligned by a magnetic field. Adv. Mater. **2002**, *14* (19), 1380–1383.
114. Haggenmueller, R.; Gommans, H.H.; Rinzler, A.G.; Fischer, J.E.; Winey, K.I. Aligned single-wall carbon nanotubes in composites by melt processing methods. Chem. Phys. Lett. **2000**, *330*, 219–225.
115. Sandler, J.K.W.; Kirk, J.E.; Kinloch, I.A.; Shaffer, M.S.P.; Windle, A.H. Ultra-low electrical percolation threshold in carbon-nanotube-epoxy composites. Polymer **2003**, *19* (44), 5893–5899.
116. Benoit, J.-M.; Corraze, B.; Lefrant, S.; Bernier, P.; Chauvet, O. Electric Transport Properties and Percolation in Carbon Nanotubes/PMMA Composites. In *Making Functional Materials with Nanotubes*; Mater. Res. Soc. Symp. Proc., Materials Research Society, 2002; Vol. 706, 85–90.
117. Coleman, J.; Dalton, A.; Curran, S.; Rubio, A.; Davey, A.; Drury, A.; McCarthy, B.; Lahr, B.; Ajayan, P.; Roth, S.; Barklie, R.; Blau, W. Phase separation of carbon nanotubes and turbostratic graphite using a functional organic polymer. Adv. Mater. **2000**, *12*, 213–216.
118. Grimes, C.A.; Mungle, C.; Kouzoudis, D.; Fang, S.; Eklund, P.C. The 500 MHz to 5.50 GHz complex permittivity spectra of single-wall carbon nanotube-loaded polymer composites. Chem. Phys. Lett. **2000**, *319* (5,6), 460–464.
119. Ago, H.; Shaffer, M.S.P.; Ginger, D.S.; Windle, A.H.; Friend, R.H. Electronic interaction between photoexcited poly(p-phenylene-vinylene) and carbon nanotubes. Phys. Rev., B **2000**, *61*, 2286.
120. Woo, H.S.; Czerw, R.; Webster, S.; Carroll, D.L.; Ballato, J.; Strevens, A.E.; O'Brien, D.; Blau, W.J. Hole blocking in carbon nanotube-polymer composite organic light emitting diodes based on poly(m-phenylene-vinylene-co-2,5 dioctoxy-p-phenylene). Appl. Phys. Lett. **2000**, *77*, 1393–1395.
121. Tang, B.Z.; Xu, H. Preparation, alignment, and optical properties of soluble poly(phenylacetylene)-wrapped carbon nanotubes. Macromolecules **1999**, *32*, 2569.
122. Nielsen, L.; Landel, R. *Mechanical Properties of Polymers and Composites*; Marcel Dekker Inc.: New York, 1994. Chapter 7.
123. Odegard, G.M.; Gates, T.S.; Wise, K.E.; Park, C.; Siochi, E.J. Constitutive modeling of nanotube-reinforced polymer composites. Compos. Sci. Technol. **2003**, *63*, 1671–1687.
124. Ji, X.L.; Jing, J.K.; Jiang, W.; Jiang, B.Z. Tensile modulus of polymer nanocomposites. Polym. Eng. Sci. **2002**, *42*, 983–993.

Polymer Nanofibers Prepared by Electrospinning

Roland Dersch
Andreas Greiner
Joachim H. Wendorff
Philipps-Universität Marburg, Marburg, Germany

INTRODUCTION

The use of fibers has a long history, the traditional application being in clothing. Natural fibers such as silk, wool, or cotton provide attractive properties for such applications. Fiber diameters characteristic of such natural fibers are 15–40 μm for sheep wool, approximately 20 μm for cotton, and approximately 20 μm for cultivated silk. Natural fibers compete in this area of applications with synthetic fibers. Man-made fibers are predominantly produced by melt spinning, which leads to fiber diameters typically in the range from about 5 to more than 70 μm, thus covering the diameter range characteristic of natural fibers. To produce such fibers, the polymer melt is pumped through a spinneret. Fiber stretching is applied not only to obtain a certain reduction of the diameter but also to simultaneously enhance the chain orientation and mechanical properties.

OVERVIEW

The range of fiber end-uses is obviously not restricted to clothing. Among others, fibers are used as sorbants, for thermal insulation, filtration, protective clothing, in composites to reinforce a polymer matrices, etc. Such applications frequently require the reduction of the fiber diameter to obtain microfibers, ultramicrofibers down to nanofibers, for example for increase of efficiency by increase of the specific surface area. Polymer fibers with diameters down to 10 nm and below are in demand for a broad range of applications to be discussed below in more detail. These include nanoscaled reinforcement, tissue engineering, specialty fibers, and templates for the formation of hollow fibers with inner diameters in the nanometer range.[1–9]

Melt blowing invented close to 50 years ago has become a process of choice for the production of microfibers. It is basically a process in which a high-velocity stream of gases or fluids—steam, air, or other fluids—blows molten thermoplastic resins from an extruder die tip onto a substrate, a conveyor, or a take-up screen. The characteristic feature is that this process provides a fine fiber web; that is, nonwovens rather than textile fibers.

Another approach toward microfibers involves the spinning of two polymers together to form fine-diameter fibers within the matrix of the second polymer. The matrix is dissolved to yield the microfiber. A related approach consists in splitting the composite fiber by mechanical forces, a high-velocity stream of water. None of these techniques is able to produce polymer fibers down to the nanometer range. The only technique currently available to yield such extremely thin fibers is electrospinning, which is also called electrostatic spinning.

The intention of this review on electrospinning of polymers is not to be comprehensive nor to demonstrate the historical evolution but rather to show concepts and disclose further challenges in electrospinning.

ELECTROSPINNING PROCESS: THEORETICAL BACKGROUND

The experimental observation is that electrospinning is able to yield well-defined nanofibers with diameters down to a few nanometers for many polymers. Yet a broad range of parameters seems to influence fiber formation and it is a frequent observation that electrospinning does not give rise to homogeneous fibers but rather to beaded fibers; that is, fibers in combination with droplets (Fig. 1). Theory is helpful for understanding the main mechanism causing fiber formation and how particular parameters affect fiber formation.

In electrospinning, a strong electrical field is applied to a droplet formed by a polymer solution or polymer melt at the tip of a die acting as one of the electrodes (Fig. 2). The charging of the fluid leads, as shown by Taylor[10] in a set of papers, to a conical deformation of the droplet—the well-known Taylor cone and eventually to the ejection of a jet from the tip of the cone. Taylor found that the critical field required for such deformation is strongly controlled by the surface tension. The electric dispersion of liquids based on this phenomenon is widely used today, for instance in ink jet printing, for crop spraying, to prepare aerosols from liquids, and to produce propellants for rockets.

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013605

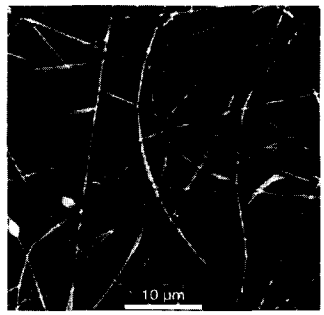


Fig. 1 Beaded poly(vinylalcohol) fibers by electrospinning from aqueous solution.

In electrospinning of fibers, the conventional view is that the charged viscoelastic jet is accelerated toward the counter electrodes, and rapidly thins during this period because of elongation and evaporation of the solvent. In some cases, splitting of the jet into multiple jets has been observed leading to a further thinning.[11] Solidification takes place and solid fibers are randomly deposited onto the substrate with a deposition rate in the order of several meters per second.[12]

However, a more detailed experimental and theoretical analysis has shown that the electrospinning process is much more complex.[13–19] It is controlled by various types of instabilities such as the Rayleigh instability, an axisymmetric instability, and finally by a so-called whipping or bending instability. In particular, the whipping instability has been identified as the one controlling elongation and thinning of the electrospun fibers. The other types of instability causes fluctuations of the radius of the jet and may eventually result in droplet formation.

The Rayleigh instability in the presence of electrical fields is the electrical counterpart of the well-known surface energy driven Rayleigh instability observed for liquid threads in the absence of external forces. This type of instability is characterized by statistical fluctuation of the radius of the jet and the growth of such fluctuations with specific wavelength. The effect of an applied field is that the wavelength for fluctuations with positive growth rate is shifted to larger and larger wavelength with increasing field strength until these fluctuations are finally suppressed at a still larger field strength.

However, the jet is still unstable because a new type of instability—the so-called axisymmetric mode. A statistical modulation of the radius of the jet induces a modulation of the density of the surface charges. These modulations result in tangential stresses, which in turn accelerate the liquid and thus enhance the modulation of the radius. Again, the formation of droplets can be the final state. It is the absolute value of the charge density at the surface, which controls other factors when this instability becomes dominant.

Finally, the whipping or bending type of instability occurs at higher field strengths and this is apparently the most important type for electrospinning of fibers. The strong thinning of the jet as well as the strong elongational deformation, which takes place during electrospinning, has been attributed to this mode of instability. For large static charge densities, the whipping mode tends to dominate because high-charge densities at the surface simultaneously tend to suppress both the Rayleigh and the axisymmetric mode of instability. It has even been claimed that the onset threshold for electrospinning corresponds to the excitation of the whipping mode.

The whipping mode corresponds to long-wavelength oscillations of the centerline of the jet: that is, the jet is subjected to bending modes. Dipolar components arise if a charged jet is subjected to bending because the charge density is no longer uniform across the cross section of the jet. An elastic force resulting from the viscous flow of the liquid counteracts the bending mode. This force is related to the rate of strain and thus to the time derivative of the curvature of the jet. For larger static charge densities, the whipping mode tends to dominate because high surface charges simultaneously suppress the axisymmetric mode.

These theoretical considerations point out the important role of surface charge and thus of the conductivity of the material used for spinning as well as of the surface energy. The role of viscoelastic properties of solutions or melts are less well known and require further experimental investigations. Furthermore, the expectations are that

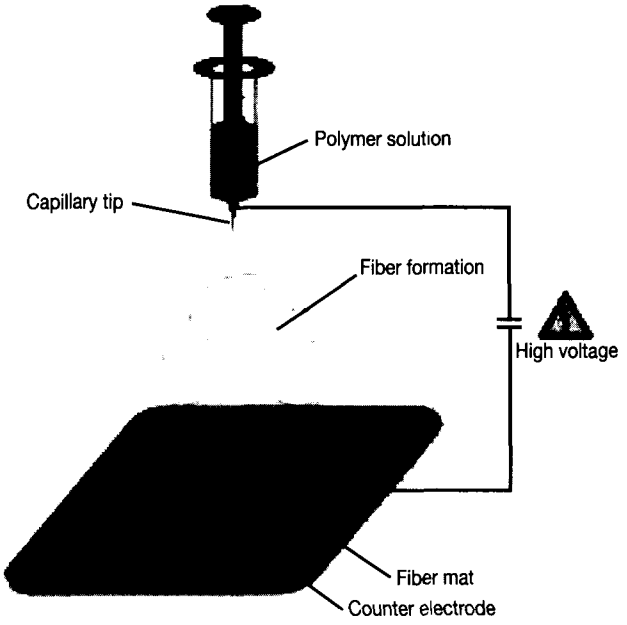


Fig. 2 Schematical graph of the electrospinning setup. (*View this art in color at www.dekker.com.*)

structure formation processes have to rapidly occur during electrospinning because of the large deposition rates, which in turn should lead to nonequilibrium structures within the nanofibers. A second feature should be high orientations as a result of the strong deformation, which occur in the whipping mode.

PARAMETERS AFFECTING FIBER FORMATION

Fiber diameters and the absence or presence of beads are features that have to be reproducibly controlled. A set of parameters including surface free energy, viscosity, electrical conductivity, molecular weight, and its distribution control fiber formation can have an effect on the formation of beaded fibers. In the following, experimental results will be discussed exploring the importance of these parameters.

Surface Free Energy

The critical field for the onset of fiber formation is predicted to be proportional to the square root of the surface tension. This finding is roughly in agreement with experimental findings. The more important question is whether this parameter affects fiber diameters. The observation is that a variation of surface free energy keeping the other controlling parameters constant did not significantly modify the fiber diameter.

Electrical Conductivity

A second key parameter is the conductivity of the melt or solution.[20,21] An increase of the conductivity is expected to increase the surface charge density, which in turn should lead to a stabilization of the whipping modes giving rise to fiber extension. Furthermore, bead formation arising from other modes of instability should be suppressed. Experiments performed on solutions of polylactide (PLA) in dichloromethane doped with various amount of pyridinium formate showed that in fact an increasing concentration of the dopant lead to a reduction of bead formation and to smaller diameters of the fibers (Fig. 3).

These results agree in principle with those from Baumgarten.[22] The prediction was that the final fiber diameter depends on the conductivity and that it varies as

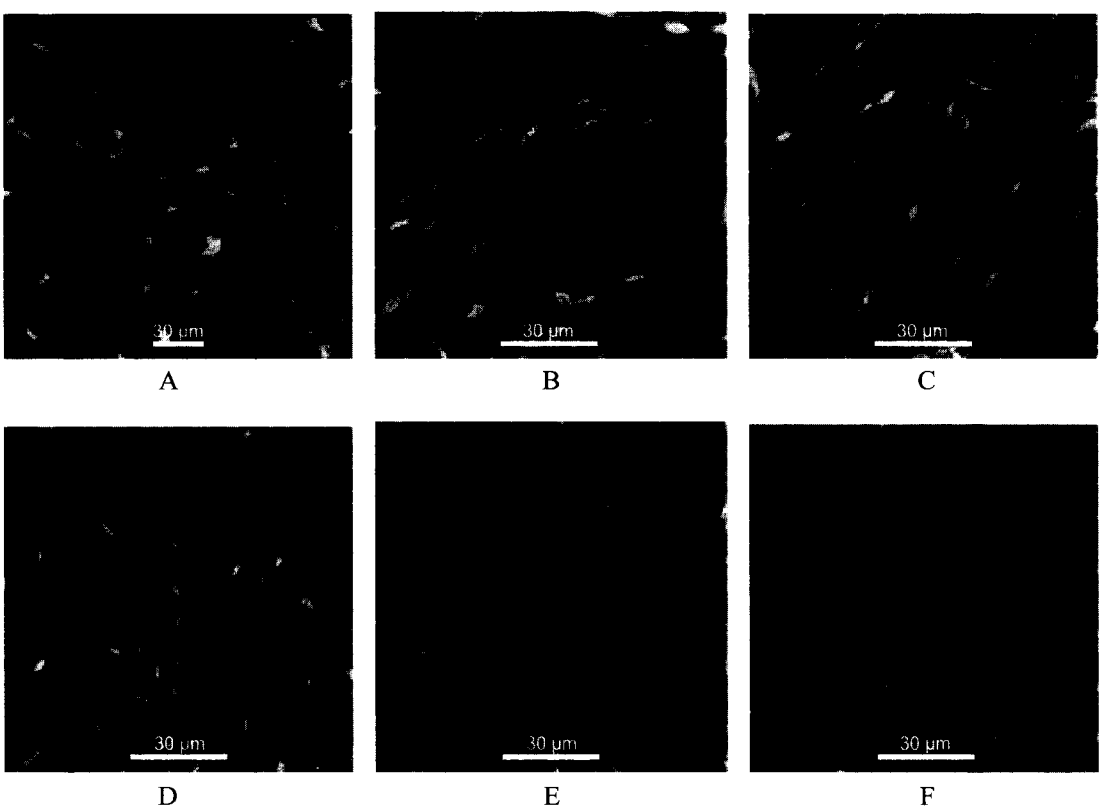


Fig. 3 Scanning electron microscopy images of electrospun PLA fibers from 2% PLA solution in dichloromethane with different concentrations (w/w in solution) of pyridinium formate: (A) 0%; (B) 0.2%; (C) 0.3%; (D) 0.4%; (E) 0.6%; (F) 0.8%.

the cube root of the resistivity. Baumgartner also stated that only droplets are generated if the conductivity is too high and he also reported that no continuous spinning could be achieved if the conductivity is too low. The reason is that the flow rate is too small in this case.

Concentration of the Polymer in Solution

There is a direct correlation between concentration and fiber diameter. For instance, a decrease of the PLA concentration in dichloromethane from 5 to 0.8 wt.% causes a reduction of the fibers from approximately 500–1000 to approximately 10–70 nm.[20] Provisions have to be made to reduce bead formation. Electrospinning from more diluted solutions requires a certain degree of electrical conductivity.

Subsequent investigations revealed that the control of the fiber diameter, the presence or absence of drop formation, the control of the surface morphology, and also the control of the texture of the web that is produced by electrospinning depend on a much broader range of parameters. These are among other thermodynamic properties of the solvent and the polymer—vapor pressure, crystallization and glass transition temperature, solubility of the polymer in the solvent or mixture of solvent. A reduction in molecular weight or an increase of the concentration of the lower molecular weight component in blends have been found to cause a reduction of the fiber diameter.

The concept of electrospinning is simple; yet, it is a demanding task to adjust the electrospinning parameters for a given system to be able to spin nanofibers in a controlled manner.

A BRIEF SURVEY ON ELECTROSPUN POLYMER SYSTEMS

A broad range of polymer systems has been subjected to electrospinning. Most of the papers published on electrospinning of polymer fibers consider spinning from solution. One advantage of electrospinning is that water can be used as a solvent. Water-soluble polymers such as poly(ethylenoxide) (PEO) or poly(vinylalcohol) (PVA) can thus be electrospun.[21,23,24] Electrospinning without any solvent from the melt was reported by Larrando and Manley[25–27] for poly(propylene) and poly(ethylene terephthalate), although true nanofibers as major products have not been obtained so far by melt electrospinning. Besides PEO and PVA, electrospinning of numerous polymers to nanofibers has been reported including poly(acrylonitrile),[28] PLA,[9,20,29] poly(styrene),[11] poly (methylmethacrylate),[30] poly(amides),[12] poly (imides),[31] poly(caprolactone),[32,33] and poly(vinylidene fluoride),[11] just to name a few.

Electrospinning of blends[34–37] or composites[9,37–41] considerably broadens the range of fiber properties and applications and might lead to novel functionalities. Polymer nanofibers can be loaded with drugs or other biologically active molecules by spinning of polymers from ternary solutions containing the reagent. For example, Kenawy et al.[42] studied the release of tetracycline hydrochloride from poly(ethylene-co-vinyl-acetate), PLA, and a blend thereof. Similarly, but with a different purpose, He[43] investigated electrospinning of PVA–Pt/TiO_2 and subsequent photocatalytic degradation of the PVA nanofibers. Core–shell fibers with concentric variation of concentrations were obtained by concentric coelectrospinning of a solution of palladium(II)acetate (core) and PEO (shell),[44] which can be extended to other systems as well.

STRUCTURE FORMATION DURING ELECTROSPINNING

A characteristic feature of the electrospinning process is the rapid evaporation of the solvent in the case of solution spinning and the rapid temperature decrease in the case of electrospinning from the melt. Thus structure formation has to happen on a millisecond scale. The nucleation of crystals should therefore be strongly quenched. A second feature is the strong deformation taking place during the whipping mode, which should give rise to orientational process within the fibers.

Experiments performed on nanofibers made from polymers able to crystallize such as polyamides, polyethylene oxide, or PLA have revealed that crystals are able to grow. In fact, the degree of crystallinity and the perfection of the crystal are not much different from those observed for thicker fibers obtained from melt extrusion or for thicker films.[45]

High degrees of crystal orientation have been found for various electrospun polymers including polyethylene oxide and polyamides (Fig. 4). Performing selected area electron diffraction studies, it was shown for polyamide fibers that the degree of orientation directly obtained by electrospinning corresponds to the one obtained for melt-extruded fibers only after considerable stretching. On the other hand, no crystal orientation has been reported for electrospun fibers made from PLAs.[46] It is currently not obvious why crystal orientation is absent in this case. Crystal formation, degree of crystallinity, and the crystal orientation are parameters that considerably affect properties and functions of the fibers.

The properties and functionality of electrospun nanofibers can be also modified by their surface morphology.

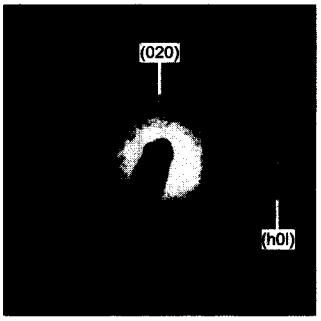


Fig. 4 Crystal orientation in PA6 fibers as obtained from selective area electron diffraction on an individual fiber with a diameter of 50 nm.

The concept is to use phase-separation processes during electrospinning. One approach considers electrospinning of ternary systems composed of two incompatible polymers and a solvent.[34] The resulting nanofibers are characterized in this case by phase morphologies corresponding either to cocontinuous spinodal type morphologies or dispersed binodal-like morphologies. Such internally structured fibers are of interest for many applications. Yet one might also selectively remove one of the components to generate fibers characterized by porous structures or other types of topologies. Such surface topologies are known to affect surface free energy, wetting, as well as adhesion. The experimental finding is that such types of phase separation processes may also occur during electrospinning in binary systems composed of one polymer and a solvent.[29,30] PLA nanofibers electrospun from dichloromethane solution typically show a nanoporous surface (Fig. 5).

Nanoporous electrospun fibers were also obtained for a variety of other polymer systems such as polycarbonates or polystyrene.[30] Interestingly, electrospinning at different humidities had a significant effect on the porosity of the electrospun fibers. It has been concluded that pore formation arising from rapid solvent evaporation most likely occurred because of the formation of water droplets from atmospheric water as a result of the evaporative cooling of the polymer solution on its way to the counter electrode in the electrospinning process.

APPLICATIONS OF ELECTROSPUN NANOFIBERS

Electrospun nanofibers offer wide applications such as in the fields of specialty filters,[47] tissue engineering,[4] wound healing,[48] reinforcement,[1,49] protective clothing,[2,50] drug release,[42] sensors,[51] catalysis,[52,53] and templating,[8,9,54,55] to name a few. Presently, most patents on electrospinning are focused on applications rather than on new materials or on new processes. Commercialized applications were accomplished by several companies worldwide in specialty filters (air filtration, coalescence filters) with an increasing market in the coming years. Tissue engineering is going to be another successful field of applications for electrospun biodegradable nanofibers (e.g., PLA) because cells including stem cells grow well on these nanofibers. An additional plus is the large potential for loading of electrospun nanofibers by biologically active agents ranging from salts to drugs and proteins.[42,56] Protective clothing, in particular in combination with special reagents, is expected to enhance the performance of military protective clothing.[6] The main focus here is on trapping of aerosols of chemical and biological warfare agents.

A unique application of electrospun nanofibers is their use as templates for the preparation of other nano-objects such as nanotubes.[8,9,55,56] Coating of electrospun nanofibers by different materials and coating techniques and subsequent removal of the core template fibers generates a whole set of nanotubes of different sizes and materials

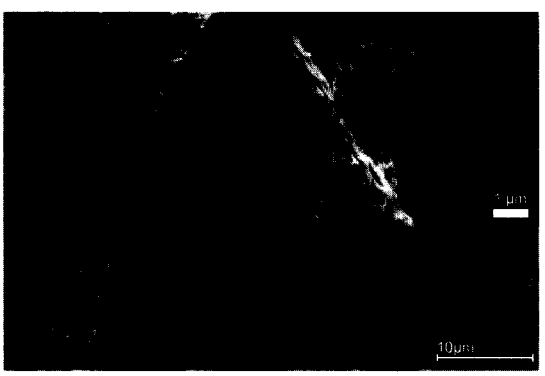


Fig. 5 Nanoporous electrospun PLA nanofibers from a dichloromethane solution.

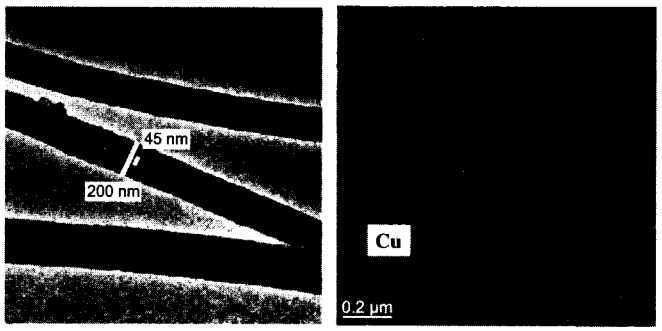


Fig. 6 Transmission electron micrograph of poly(*p*-xylylene) (PPX) nanotubes (A) and of PPX nanotubes loaded by copper nanoparticles.

combinations, which has by far not been exploited (Fig. 6). The versatility of this approach is enhanced again by the option of loading of electrospun nanofibers directly during electrospinning or coelectrospinning.[44]

CONCLUSION

As electrospinning of polymers is a rather old technique that has just recently been rediscovered, a new community has been established dealing with electrospinning. The number of papers is exponentially increasing as obvious from the papers published per year. The technique will be certainly optimized within the next few years but a precise understanding of the impact of different processing parameters based on empirical observations together with theoretical considerations will be a key issue for optimum results. Many new open questions remain concerning materials choice but major new developments are expected in the area of structuring of electrospun nanofibers, loading of nanofibers by new additives for special applications, and the development of smart nanofibers. Clearly, a major driving force of new developments in the field of electrospun nanofibers will be new applications, which nicely unites researchers from different areas such as polymers, physics, engineering, inorganic and organic chemistry, medicine, pharmacy, or rheology, to name a few.

REFERENCES

1. Bergshoef, M.M.; Vancso, G.J. Transparent nanocomposites with ultrathin, electrospun nylon-4,6 fiber reinforcement. Adv. Mater. **1999**, *11*, 1362.
2. Gibson, P.W.; Schreuder-Gibson, H.L.; Rivin, D. Electrospun fiber mats: Transport properties. AIChE J. **1999**, *45*, 190.
3. Boland, E.; Wnek, G.; Simpson, D.; Pawlowski, K.; Bowlin, G.L. Tailoring tissue engineering scaffolds using electrostatic processing techniques: A study of poly(glycolic acid) electrospinning. J. Macromol. Sci., Pure Appl. Chem. **2001**, *38*, 1231.
4. Li, W.J.; Laurencin, C.T.; Caterson, E.J.; Tuan, R.S.; Ko, F.K. Electrospun nanofibrous structure: A novel scaffold for tissue engineering. J. Biomed. Mater. Res. **2002**, *60*, 613.
5. Shortkroff, S.; Li, Y.; Thornhill, T.S.; Rutledge, G.C. Cell growth on electrospun PCL scaffolds. Polym. Mater. Sci. Eng. **2002**, *87*. 457.
6. Schreuder-Gibson, H.; Gibson, P.; Senecal, K.; Sennett, M.; Walker, J.; Yeomans, W.; Ziegler, D.; Tsai, P.P. Protective textile materials based on electrospun nanofibers. J. Adv. Mater. **2002**, *34*, 44.
7. Tsai, P.P.; Schreuder-Gibson, H.; Gibson, P. Different electrostatic methods for making electret filters. J. Electrost. **2002**, *54*, 333.
8. Bognitzki, M.; Hou, H.Q.; Ishaque, M.; Frese, T.; Hellwig, M.; Schwarte, C.; Schaper, A.; Wendorff, J.H.; Greiner, A. Polymer, metal, and hybrid nano- and mesotubes by coating degradable polymer templates fibers (TUFT process). Adv. Mater. **2000**, *12*, 637.
9. Hou, H.; Jun, Z.; Reuning, A.; Schaper, A.; Wendorff, J.H.; Greiner, A. Poly(*p*-xylylene) nanotubes by coating and removal of ultrathin polymer template fibers. Macromolecules **2002**, *35*, 2429.
10. Taylor, G.I. Electrically driven jets. Proc. R. Soc. A **1969**, *313*, 453.
11. Koombhongse, S.; Liu, W.X.; Reneker, D.H. Flat polymer ribbons and other shapes by electrospinning. J. Polym. Sci., Part B, Polym. Phys. **2001**, *39*, 2598.

12. Reneker, D.H.; Chun, I. Nanometre diameter fibres of polymer, produced by electrospinning. Nanotechnology **1996**, *7*, 216.
13. Hohman, M.M.; Shin, M.; Rutledge, G.; Brenner, M.P. Electrospinning and electrically forced jets. I. Stability theory. Phys. Fluids **2001**, *13*, 2201.
14. Hohman, M.M.; Shin, M.; Rutledge, G.; Brenner, M.P. Electrospinning and electrically forced jets. II. Applications. Phys. Fluids **2001**, *13*, 2221.
15. Yarin, A.L.; Koombhongse, S.; Reneker, D.H. Bending instability in electrospining of nanofibers. J. Appl. Phys. **2001**, *89*, 3018.
16. Shin, Y.M.; Hohman, M.M.; Brenner, M.P.; Rutledge, G.C. Electrospinning: A whipping fluid jet generates submicron polymer fibers. Appl. Phys. Lett. **2001**, *78*, 1149.
17. Reneker, D.H.; Yarin, A.L.; Fong, H.; Koombhongse, S. Bending instability of electrically charged liquid jets of polymer solutions in electrospinning. J. Appl. Phys. **2000**, *87*, 4531.
18. Yarin, A.L.; Koombhongse, S.; Reneker, D.H. Taylor cone and jetting from liquid droplets in electrospinning of nanofibers. J. Appl. Phys. **2001**, *90*, 4836.
19. Fridrikh, S.; Yu, J.; Brenner, M.; Rutledge, G.C. Controlling the fiber diameter during electrospinning. Phys. Rev. Lett. **2003**, *90*. art. no. 144502.
20. Jun, Z.; Hou, H.; Schaper, A.; Wendorff, J.; Greiner, A. Poly-L-lactide nanofibers by electrospinning—Influence of solution viscosity and electrical conductivity on fiber diameter and fiber morphology. E-Polymers **2003**. art. no. 009.
21. Doshi, J.; Reneker, D.H. Electrospinning process and applications of electrospun fibers. J. Electrost. **1995**, *35*, 151.
22. Baumgarten, P. Electrostatic spinning of acrylic microfibers. J. Colloid Interface Sci. **1971**, *36*, 71.
23. Ding, B.; Kim, H.Y.; Lee, S.C.; Shao, C.L.; Lee, D.R.; Park, S.J.; Kwag, G.B.; Choi, K.J. Preparation and characterization of a nanoscale poly(vinyl alcohol) fiber aggregate produced by an electrospinning method. J. Polym. Sci., Part B, Polym. Phys. **2002**, *40*, 1261.
24. Jaeger, R.; Bergshoef, M.M.; Batlle, C.M.I.; Schönherr, H.; Vancso, G.J. Electrospinning of ultra-thin polymer fibers. Macromol. Symp. **1998**, *127*, 141.
25. Larrondo, L.; Manley, R.S.J. Electrostatic fiber spinning from polymer melts 2. Examination of the flow field in an electrically driven jet. J. Polym. Sci., Part B, Polym. Phys. **1981**, *19*, 921.
26. Larrondo, L.; Manley, R.S.J. Electrostatic fiber spinning from polymer melts 3. Electrostatic deformation of a pendant drop of polymer melt. J. Polym. Sci., Part B, Polym. Phys. **1981**, *19*, 933.
27. Larrondo, L.; Manley, R.S.J. Electrostatic fiber spinning from polymer melts 1. Experimental—Observations on fiber formation and properties. J. Polym. Sci., Part B, Polym. Phys. **1981**, *19*, 909.
28. Chun, I.; Reneker, D.; Fong, H.; Fang, X.; Dietzel, J.; Tan, N.; Kearns, K. Carbon nanofibers from polyacrylonitrile and mesophase pitch. J. Adv. Mater. **1999**, *31*, 36.
29. Bognitzki, M.; Czado, W.; Frese, T.; Schaper, A.; Hellwig, M.; Steinhart, M.; Greiner, A.; Wendorff, J.H. Nanostructured fibers via electrospinning. Adv. Mater. **2001**, *13*, 70.
30. Megelski, S.; Stephens, J.; Chase, D.; Rabolt, J.F. Micro- and nanostructured surface morphology on electrospun polymer fibers. Macromolecules **2002**, *35*, 8456.
31. Nah, C.; Han, S.; Lee, M.; Kim, J.; Lee, D.S. Characteristics of polyimide ultrafine fibers prepared through electrospinning. Polym. Int. **2003**, *52*, 429.
32. Reneker, D.H.; Kataphinan, W.; Theron, A.; Zussman, E.; Yarin, A.L. Nanofiber garlands of polycaprolactone by electrospinning. Polymer **2002**, *43*, 6785.
33. Lee, K.; Kim, H.; Khil, M.; Ra, Y.; Lee, D.R. Characterization of nano-structured poly(epsilon-caprolactone) nonwoven mats via electrospinning. Polymer **2003**, *44*, 1287.
34. Bognitzki, M.; Frese, T.; Steinhart, M.; Greiner, A.; Wendorff, J.H.; Schaper, A.; Hellwig, M. Preparation of fibers with nanoscaled morphologies: Electrospinning of polymer blends. Polym. Eng. Sci. **2001**, *41*, 982.
35. Huang, L.; Nagapudi, K.; Apkarian, R.; Chaikof, E.L. Engineered collagen–PEO nanofibers and fabrics. J. Biomater. Sci., Polym. Ed. **2001**, *12*, 979.
36. Lee, S.H.; Yoon, J.W.; Suh, M.H. Continuous nanofibers manufactured by electrospinning technique. Macromol. Res. **2002**, *10*, 282.
37. Wu, G.; Miura, T.; Asai, S.; Sumita, M. Carbon black-loading induced phase fluctuations in PVDF/PMMA miscible blends: Dynamic percolation measurements. Polymer **2001**, *42*, 3271.
38. Kenawy, E.R.; Bowlin, G.L.; Mansfield, K.; Layman, J.; Simpson, D.G.; Sanders, E.H.; Wnek, G.E. Release of tetracycline hydrochloride from electrospun poly(ethylene-co-vinylacetate), poly(lactic acid) and a blend. J. Control. Release **2002**, *81*, 57.
39. Wang, C.; Hong, Y.; Qingbiao, Y.; Yunus, T.; Li, Z.; Guo, N.; Zhao, Y.; Liu, W.; Hao, X.; Fei, Q. Array of nanoparticles in PAN nanofiber. Polym. Prep. **2002**, *43*, 1387.
40. Yang, Q.; Li, D.; Hong, Y.; Li, Z.; Wang, C.; Qiu, S.; Wei, Y. Preparation and characterization of a pan

nanofibre containing Ag nanoparticles via electrospinning. Synth. Met. **2003**, *137*, 973.
41. Fong, H.; Liu, W.D.; Wang, C.S.; Vaia, R.A. Generation of electrospun fibers of nylon 6 and nylon 6-montmorillonite nanocomposite. Polymer **2002**, *43*, 775.
42. Kenawy, E.; Bowlin, G.; Mansfield, K.; Layman, J.; Simpson, D.; Sanders, E.; Wnek, G.E. J. Control. Release **2002**, *81*, 57.
43. He, C.G.J. The preparation of PVA–Pt/TiO_2 composite nanofiber aggregate and the photocatalytic degradation of solid-phase polyvinyl alcohol. Polym. Degrad. Stab. **2003**, *81*, 117.
44. Zussman, E.; Yarin, A.; Sun, Z.; Wendorff, J.; Greiner, A. Compound core/shell polymer nanofibers by co-electrospinning. Adv. Mater. **2003**, *15*, 1929.
45. Dersch, R.; Liu, T.; Schaper, A.; Greiner, A.; Wendorff, J.H. Electrospun nanofibers: Internal structure and intrinsic orientation. J. Polym. Sci., Polym. Chem. **2003**, *41*, 545.
46. Zong, X.; Kim, K.; Fang, D.; Ran, S.; Hsiao, B.; Chu, B. Structure and process relationship of electrospun bioabsorbable nanofiber membranes. Polymer **2002**, *43*, 4403.
47. Jacobsen, M. The nonwovens industry meets the filtration business. Nonwovens Ind. May **1991**, 36.
48. Jin, H.; Fridrikh, S.; Rutledge, G.; Kaplan, D.L. Electrospinning Bombyx mori silk with poly(ethylene oxide). Biomacromolecules **2002**, *3*, 1233.
49. Kim, J.S.; Reneker, D.H. Mechanical properties of composites using ultrafine electrospun fibers. Polym. Compos. **1999**, *20*, 124.
50. Schreuder-Gibson, H.; Gibson, P.; Senecal, K.; Sennett, M.; Walker, J.; Yeomans, W.; Ziegler, D.; Tsai, P.P. Protective textile materials based on electrospun nanofibers. J. Adv. Mater. **2002**, *34*, 44.
51. Wang, X.; Drew, C.; Lee, S.; Senecal, K.; Kumar, J.; Samuelson, L.A. Electrospinning technology: A novel approach to sensor application. J. Macromol. Sci., Pure Appl. Chem. **2002**, *A39*, 1251.
52. Jun, Z.; Sun, Z.; Greiner, A., submitted.
53. Jia, H.; Zhu, G.; Vugrinovich, B.; Kataphinan, W.; Reneker, D.; Wang, P. Enzyme-carrying polymeric nanofibers prepared via electrospinning for use as unique biocatalysts. Biotechnol. Prog. **2002**, *18*, 1027.
54. Caruso, R.A.; Schattka, J.H.; Greiner, A. Titanium dioxide tubes from sol–gel coating of electrospun polymer fibers. Adv. Mater. **2001**, *13*, 1577.
55. Dong, H.; Jones, W. A fiber templating approach to conducting polymer nanotubes. Polym. Mater. Sci. Eng. **2002**, *87*, 273.
56. Luu, Y.; Kim, K.; Hsiao, B.; Chu, B.; Hadjiargyrou, M. Development of a nanostructured DNA delivery scaffold via electrospinning of PLGA and PLA-PEG block copolymers. J. Control. Release **2003**, *89*, 341.

Polymer Nanoparticles for Gene Delivery: Synthesis and Processing

P

Jie Wen
Kam W. Leong
The Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.

INTRODUCTION

Gene therapy, a therapeutic modality designed to introduce an extraneous gene into the patient's cells to produce a specific protein, potentially can treat a variety of inherited and acquired diseases. Gene medicines are composed of a gene expression vector that encodes a therapeutic protein and a gene delivery system that delivers the gene to the desired tissue and subcellular compartments. The specific and efficient delivery of DNA to the diseased sites and to the nuclei of particular cell populations is the key for successful gene therapy. Gene delivery system plays a central role in fulfilling this mission.

OVERVIEW

Traditionally, DNA delivery systems fall into two categories, viral-mediated and nonviral-mediated. In addition, naked DNA can be directly injected into the target tissue to produce the transgenes. Viral vectors such as retrovirus, adenovirus, adeno-associated virus, lentil virus, and herpes-simplex virus are usually much more efficient than nonviral vectors in both gene delivery and expression because of their highly evolved and specialized mechanisms in infecting cells. However, viral vectors have inherent disadvantages such as immunogenicity, restricted targeting of specific cell types, size limit of the DNA packaged into a virus, and scale-up difficulties.[1] The safety issue takes on new significance in light of the tragic event in recent clinical trial.[2] Consequently, nonviral delivery systems have been increasingly proposed as alternatives to viral systems. However, nonviral systems are inefficient and produce only transient transgene expression.

Nonviral gene delivery systems rely on condensing DNA to a compact size that can be internalized by cell and protecting the DNA from enzymatic degradation until the DNA reaches the nucleus. However, this tight binding and protection must be balanced by the ability of the gene carrier to release the DNA. The nonviral gene carriers include cationic liposomes, cationic polymers, and neutral polymers. They are able to condense DNA to nanoparticles with a size of 20–500 nm in diameter for cellular internalization. The term nanoparticle is used in this review to cover particulate carrier systems that are less than 1 μm in size and normally below 500 nm, including liposomes and solid nanocapsules.

Most cationic liposome formulations are composed of a cationic lipid and a neutral helper lipid. The cationic lipid binds the DNA through electrostatic interaction to form a complex termed lipoplex, while the helper lipid helps to stabilize the liposome and improve the cytoplasmic delivery of DNA.[3] Cationic polymers are usually nitrogen-rich polymers and have been used in much the same way as cationic liposomes with the polymer interacting electrostatically with DNA to yield a transfection-competent complex, termed polyplex. Polyplex is generally more stable than lipoplex in a physiological environment. The neutral polymer system usually involves encapsulating DNA in nanoparticles using biodegradable polymers, such as PLA and PLGA, to obtain sustained release characteristics. Recently, an inorganic system based on calcium phosphate nanoparticle has also been reported for gene delivery.[4]

Although most nonviral gene carriers can efficiently condense and protect the DNA, delivering the intact DNA to the nucleus remains a major challenge in the field. Poor mechanistic understanding of the whole gene transfer process makes it difficult to develop an optimal delivery system with all the desirable features for gene delivery. This review focuses on the application of cationic polymer-based nanoparticles for gene delivery.

NANOPARTICLE FORMATION

Because of large electrostatic repulsion, free or naked DNA remains in an expanded state. Free DNA is too large and too polar a molecule to cross cell membranes by passive diffusion. The condensation of DNA into compacted nanoparticles is essential for the transport of DNA

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009380

through the cell membrane, although recent studies suggest that naked DNA can be taken up by muscle cells via receptor-mediated endocytosis.[5] Multivalent cations such as polyamines, positively charged polymers, and peptides are known to be capable of condensing DNA via electrostatic interaction with the negatively charged phosphate groups of DNA. A valence of ≥ 3 is required for the multivalent cations to condense DNA because the Coulombic interactions should be stronger than the entropic interactions so the entropy loss can be overcome.[6] The condensed DNA nanoparticles appear as rods, toroids, or spheroids under electron microscope.[7]

In understanding the mechanism of DNA condensation, the most extensive work has been conducted using small cations, such as spermidine $[H_2N(CH_2)_3NH(CH_2)_4NH_2]$,[8] spermine $[H_2N(CH_2)_3NH(CH_2)_4NH(CH_2)_3NH_2]$,[9] and cobalt hexamine $[Co(NH_3)_6^{3+}]$.[10] A variety of techniques, such as microscopy, light scattering, fluorescence, and calorimetry, have been used. Two models, the spool model and the constant loop model, have been proposed to explain how the DNA is compacted within the toroid.[9,11] The cations instigate localized bending or distortion of DNA at a critical extent of charge neutralization, which facilitates the formation of rods and toroid-like structures.[12,13] The cations can also induce DNA condensation by rendering the DNA–solvent interactions less favorable.[14,15] Formation of a highly ordered liquid crystalline phase of DNA is believed to be involved in DNA condensation.[16]

Many factors affect the formation of DNA nanoparticles, such as chemical structure of cations, ionic strength of the medium, and molar ratio of the polycation and DNA. The study on the structural arrangement of DNA condensed by polyamine homologues indicates that the number of positive charges and the charge density of polyamine molecules affect the size of DNA nanoparticles as well as the ability to induce the formation of liquid crystalline phase of DNA.[17,18] In high ionic strength solution, DNA nanoparticles may aggregate because the salt overcomes the repulsive interaction between particles. Excess positive charge on the surface can prevent the nanoparticles from aggregating under the salt condition. The surface charge of the nanoparticles is characterized by ξ-potential and can be adjusted by changing the reactant ratio of DNA to polycations.

A better understanding of the relationship between the physicochemical properties of the DNA nanoparticles, their ability to be taken up by cells, and the chemical structure of agents that cause the condensation of DNA is needed for the development of nonviral gene delivery vehicles.

DELIVERY OBSTACLES FOR NONVIRAL SYSTEMS

Ultimately, DNA has to be delivered to the nucleus for successful gene expression, after overcoming a number of obstacles on the way (Fig. 1). These delivery obstacles are both extracellular as well as intracellular. The biodistribution of nanoparticles in the extracellular space is determined by their interaction with the physiological environment, based on their physicochemical and biological properties, such as size, electrical charge, hydrophilicity–hydrophobicity, and possible ligand conjugation.

Of all the properties, particle size is of crucial importance in determining the biodistribution.[19] Any particle aggregation during the gene delivery is detrimen-

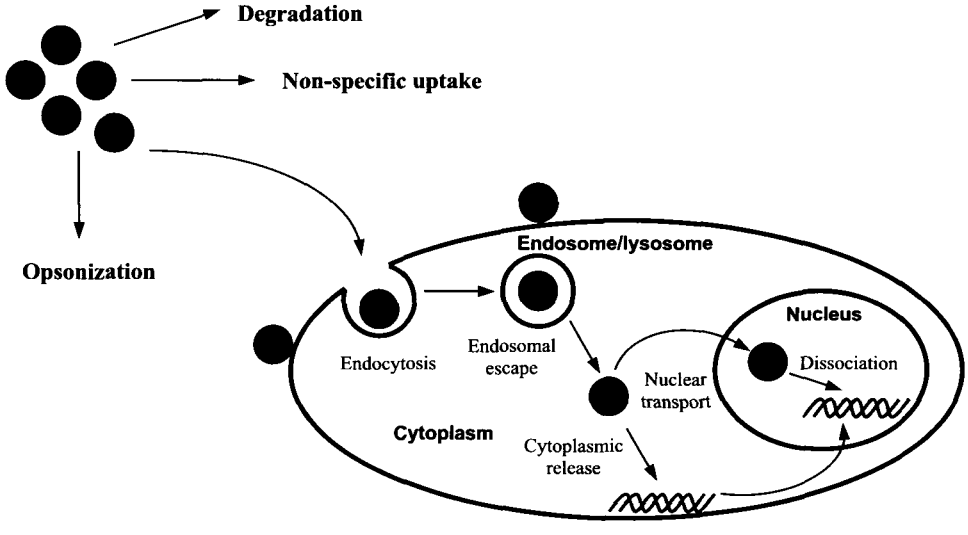


Fig. 1 Obstacles of nonviral gene delivery.

tal for DNA to get to the final site, the nucleus. In fact, negatively charged proteins such as albumin tend to bind with the positively charged nanoparticles, leading to a size increase of the particles and a reduced positive charge on the particle surface, which may diminish the affinity of the nanoparticles to the negatively charged cell surface. Therefore the nanoparticles must first be stable in the extracellular fluid so that they can be delivered to the target tissue or cells. It should be noted that the immune cells might recognize the particles as foreign materials and phagocytose them, which may be attractive for genetic vaccination but may not be desirable for other gene therapy applications.

Once the nanoparticles reach the target cells, they bind with the negatively charged cell membrane and are subsequently taken up by endocytosis. The rate of entry into the cells varies with cell type and is relatively slow.[20] In order to achieve cell-specific gene transfer, cell-specific targeting ligands such as monoclonal antibodies, peptides, and sugars can be conjugated to the gene carriers to promote receptor-mediated endocytosis. After the nonspecific or receptor-mediated endocytosis, most DNA complexes are internalized into the endosomes/lysosomes and are eventually degraded. Cytoplasmic release of the DNA complex is the first barrier in the intracellular transport. Efficient destabilization of the endosomes/lysosomes is believed to enhance gene expression. Some endosomolytic reagents such as chloroquine,[21] lipids,[22] and peptides[23] have been used to disrupt the endosome membrane, leading to an enhanced gene expression. The buffering capacity of amine-containing polymers has also proved to reduce the acidification of the endosome, triggering endosome swelling and rupture or collapse,[24] Some anionic lipids in the endosome may compete with the DNA and dissociate the DNA complexes, leading to the release of the DNA into the cytoplasm.[25]

Cytosol is not a friendly environment for DNA because of the presence of nuclease. Microinjected plasmid DNA is rapidly degraded in the cytoplasm with an apparent half-life of 50–90 min.[26] The translocation of released or "unpacked" DNA from cytosol to nucleus is probably through diffusion, a relatively slow process compared to cytoplasmic degradation. It is understandable that the DNA would be more stable in cytosol if it were still complexed with cationic carrier. However, only certain synthetic polymers, such as polyethylenimine (PEI), can show protection of DNA in the cytoplasm and also promote nuclear transport. The pathway of the transport of DNA complexes from cytoplasm to nucleus is not well understood. Some hypothesis leans toward the involvement of anionic phospholipids. The inclusion of viral nuclear localization signals (NLS) has been demonstrated to be an efficient strategy to facilitate nuclear transport.[27,28] Typically comprising several positively charged lysines and arginines, a NLS is a peptide signal sequence that directs a protein from the cytosol into the nucleus.

As stated above, eventually the DNA has to be released from the complexes to allow transcription. The dissociation of complexes may occur either in the endosome, cytosol, or nucleus. Other than enzymatic degradation of the gene carrier, anionic membrane components and other intracellular substances may also be responsible for the dissociation by carrier substitution.

POLYMERIC GENE CARRIERS

The poor transfer efficiency and transient gene expression of polymer–DNA complexes continue to stimulate the development of more effective and less-toxic polymeric gene carriers. A number of polycations have been tested to transfer genes in vitro or in vivo, including poly-L-lysine (PLL) (Fig. 2a), PEI (Fig. 2b), polyamidoamine (PAMAM) dendrimers (Fig. 2c), poly(α-(4-aminobutyl)-L-glycolic acid) (PAGA) (Fig. 2d), poly((2-dimethylamino)ethyl methacrylate) (PDMAEMA) (Fig. 2e), chitosan (Fig. 2f), etc. Our group has been interested in investigating polyphosphoester as a new polycationic gene carrier (Fig. 2g).

Poly-L-Lysine

Poly-L-lysine has been widely used in the early days as a gene delivery carrier because of its excellent DNA condensation ability and efficient protection of DNA from nuclease digestion. However, its cytotoxicity and low transfection problems remain to be solved. Moreover, PLL–DNA complexes tend to aggregate under physiological conditions. Efforts have focused on improving the solubility of PLL by the introduction of hydrophilic groups, such as dextran[29] and PEG.[30,31] Both PEG-g-PLL and PEG-b-PLL have shown lower cytotoxicity and enhanced transfection efficiency than PLL. A polymeric micelle, PLGA-g-PLL has also been proposed as an effective carrier for gene delivery.[32]

Polyethylenimine

Polyethylenimine–DNA complex is one of the most potent synthetic gene transfer vectors. They can also transfect a wide variety of cells. The high transfection efficiency of PEI has been postulated to relate to its buffering capacity, which leads to the accumulation of protons brought in by endosomal ATPase and an influx of chloride anions, triggering endosome swelling and dis-

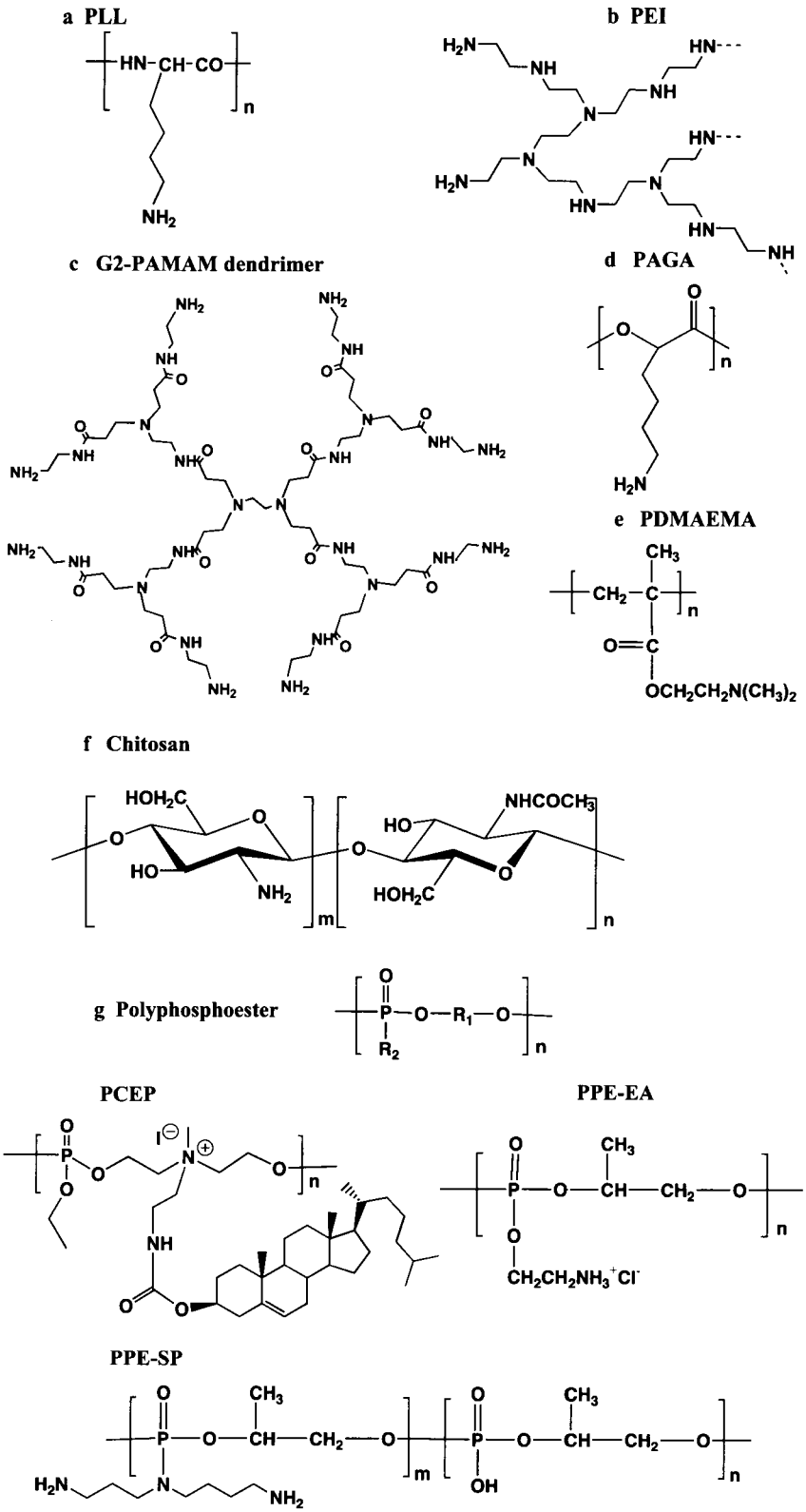


Fig. 2 Chemical structures of polymeric gene carriers.

ruption, followed by the release of DNA into cytoplasm. Significant differences were observed in the cytotoxicity and transfection efficiency of branched and linear PEI.[33] Toxicity is one of the main concerns for PEI to be used in gene delivery. Cytotoxicity of PEI is postulated to derive from its ability to permeabilize cell membranes. Moreover, PEI is not degradable and high molecular weight PEI may accumulate in the body. Several approaches have been explored to reduce the cytotoxicity, such as PEGylation[34] and conjugation of low molecular weight PEI with cleavable cross-links such as disulfide bonds, which can be cleaved in the reducing environment of the cytoplasm.[35]

Polyamidoamine Dendrimers

Starburst polyamidoamine dendrimers with either ammonia or ethylenediamine as core molecules have been used for gene delivery. These molecules are highly defined in terms of molecular weight and geometry. They are spherical polymers, with primary amine groups on the surface, which can be used to condense DNA. Dendrimers can transfect a wide variety of cells in culture,[36] with an efficiency matching that of PEI under optimal conditions. The higher generation dendrimers (G6 and G7) are effective at disrupting anionic vesicle membranes, which may explain the high transfection efficiency.[37] The advantage of these dendrimers over PEI remains to be established.

Poly(α-(4-aminobutyl)-L-glycolic acid)

Poly(α-(4-aminobutyl)-L-glycolic acid) is an analogue of PLL with a degradable ester linkage in the main chain.[38] Poly(α-(4-aminobutyl)-L-glycolic acid) efficiently condenses and protects DNA and shows enhanced transfection efficiency in culture compared to other PLL-based systems. It is also less cytotoxic than Lipofectamine and SuperFect (a dendrimers). Although PAGA generates positive results in producing cytokines as an immunoadjuvant in a murine cancer model, its biodegradation rate of the polymer may be too rapid for some in vivo gene transfer applications.[39]

Poly((2-Dimethylamino)Ethyl Methacrylate)

Poly((2-dimethylamino)ethyl methacrylate) is a water-soluble cationic polymer capable of binding plasmid DNA and transfecting different cell types.[40] The complexes with a slightly positive ξ-potential (25–30 mV) and a size around 200 nm possess the highest transfection efficiency.[41] Study on the structure–activity relationship of a number of water-soluble cationic methacrylate/methacrylamide polymers with structures closely related to PDMAEMA shows that among these polymers, PDMAEMA has the highest transfection efficiency and cytotoxicity.[42] The high transfection efficiency of PDMAEMA is probably due to its high buffering capacity to destabilize endosome, combined with its ability to release the DNA in the cytosol and/or the nucleus. Its strength and shortcomings are similar to those of PEI.

Chitosan

Chitosan, a biodegradable polysaccharide composed of D-glucosamine repeating units, has been explored by several research groups as a nonviral gene carrier.[43–45] Chitosan can efficiently bind DNA and protect DNA from nuclease degradation. Moreover, chitosan has good biocompatibility and toxicity profile, rendering it a safe biomedical material for clinical applications. Chitosan–DNA nanoparticles can transfect several different cell types. However, the transfection efficiency is relatively poor. Chitosan can be readily modified. For example, trimethylated chitosan can be prepared with different quarternization degree to increase the solubility of chitosan at neutral pH;[46] or chitosan can be conjugated with deoxycholic acid to become a colloidal gene carrier.[47] Both quaternarized chitosan and deoxylic acid-modified chitosan can efficiently transfect COS-1 cells. Chitosans with different molecular weights have different DNA binding affinities and exhibit different transfection efficiencies, indicating that particle stability may be one of the rate-limiting steps in the overall transfection process.[43] The effect of degree of deacetylation on the stability of chitosan–DNA nanoparticles is being investigated.[48]

Polyphosphoester

We have been working on polyphosphoesters (PPE) for more than 15 years, particularly the ones with a backbone analogous to nucleic acids and teichoic acids. They show promise in different biomedical applications because of their biocompatibility, biodegradability, and pendent chain functionality. Several polyphosphoesters with positive charges either in the backbone or in the side chain were synthesized and evaluated as nonviral gene carriers.[49–51] These polyphosphoesters can efficiently bind DNA and protect DNA from nuclease degradation. Efficient transfection was found in a number of cell lines, with some of them comparable to Lipofectamine. The transfection is cell-type dependant and can be improved with the incorporation of chloroquine. All the polyphosphoesters exhibit a significantly lower cytotoxicity than PLL or PEI in vitro and in vivo. With the structural versatility, polyphosphoester can be designed to have different hydophilicity–hydrophobicity properties. One interesting feature is that the PPE can provide extra-

cellular sustained release of the DNA, leading to prolonged and enhanced transgene expression in the muscle compared to naked DNA administration.[50] An amphiphilic polyphosphoester, poly{[(cholesteryl oxocarbonylamido ethyl) methyl bis(ethylene) ammonium iodide] ethyl phosphate} (PCEP), was synthesized and investigated for gene delivery.[51] Carrying a positive charge in its backbone and a lipophilic cholesterol structure in the side chain, PCEP self-assembles into micelles in aqueous solution at room temperature with an average size of 60–100 nm. This is one of the few polymeric lipids for gene delivery that are biodegradable.

TARGETING

Drug targeting can be achieved with different levels of sophistication: first-, second-, and third-order targeting. First-order targeting can be defined as the delivery of a drug to a particular organ; second-order targeting is where the drug is internalized in a specific cell type; and third-order targeting is where the drug is localized in a specific cellular compartment.[52] Gene delivery is indeed a third-order targeting challenge with nucleus as the targeting site.

The use of polymeric nanoparticles for targeted drug and gene delivery system has gained increasing attention and importance.[53] Targeting can be achieved either by passive route, where the particles are captured by a physiological uptake mechanism such as filtration or macrophage sequestration, or by active route, where the particles are recognized by the host molecule through the targeting moiety conjugated on the surface of the particles.

Passive Targeting

Normally, after the nanoparticles are injected into the blood stream, they are rapidly coated by circulation elements, such as plasma proteins and glycoproteins. This process renders the particles easily recognized by the major defense system of the body, the reticuloendothelial system. The macrophage (Kupffer) cells of the liver play a critical role in removing these coated foreign particles.[54] Thus Kupffer cells specificity can be achieved by passive targeting. Particles with a more hydrophobic surface, which can be efficiently coated with plasma elements, are more desirable for this targeting. This strategy can be used to achieve gene delivery to the blood circulation.

The lymph nodes are of great interest because of their vital role in the staging of anticancer and anti-infection immune responses.[52] A subcutaneous administration of nanoparticles can be an efficient way of reaching local draining lymph nodes. Particle size and surface property are two decisive parameters for such targeting. Generally, particles with a size less than 100 nm are preferable for such targeting. The particles with intermediate hydrophilicity–hydrophobicity in nature appear to have the most lymph node accumulation. This type targeting is particularly useful for DNA vaccination applications.

Active Targeting

To increase the specificity of DNA uptake by target cells, transferrin,[55] monoclonal antibodies,[56] mannose,[57] galactose,[58] lactose,[59] folic acid,[60] low-density lipoproteins,[61] and RGD peptides[62] have been conjugated to the polymer backbone to achieve receptor-mediated endocytosis. Poly-L-lysine and PEI are popular candidates for the conjugation because of the readily available primary amine groups. After endocytosis, the DNA nanoparticles are internalized in the vesicles within the cells. Two strategies are used to help cytoplasmic release of DNA. One is to use fusogenic peptides derived from fusion-active virus to destabilize the endosomal membrane. Another approach is to use a carrier with a high buffering capacity to promote the escape of the complexes into the cytosol, with mechanisms discussed above. The next target is then the nucleus. The cut-off size of nuclear pores is 9–10 nm, which allows proteins up to 60 kDa and nucleotides up to 300 bp to enter freely.[63] However, larger molecules can still get through the nuclear pores via an active process, which can be modulated by NLS. Enhanced gene transfection efficiency has been achieved by the conjugation of NLS peptide to the gene carriers such as PLL[64] and PDMAEMA,[65] or plasmid itself.[66] If the NLS peptide is conjugated to the gene carrier, the DNA has to stay complexed before nuclear translocation takes place. It has been reported that the number of NLS peptides conjugated to DNA plays an important role in the nuclear transport. Too many NLS peptides may actually inhibit nuclear translocation.[67]

IN VIVO APPLICATION

Targeted delivery is the ultimate goal of gene therapy. Local administration so far presents a more feasible approach as systemic barriers are bypassed. Significant gene expression has been found in a variety of tissues/organs, such as muscle, lung, liver, brain, tumor, and gastrointestinal (GI) tract using polymeric gene carriers.

Muscle

Muscle has been of interest as a target for gene therapy since Wolff et al. demonstrated that intramuscularly injected naked plasmid DNA can be expressed in myofibers in 1990.[68] Naked DNA has poor bioavailability in muscle due to the degradation by extracellular

nucleases and the rapid elimination from the muscle. In order to protect DNA from degradation cationic agents are used to condense DNA. However, the use of condensed DNA particles to achieve higher transfection in muscle rather than naked DNA has not been successful. This is possibly because the connective tissue surrounding myofibers and muscle fasciculus could limit the diffusion of rigid condensed DNA particles through the muscle.[69] Recently, a sustained release system using PPE-EA as a carrier has been proved to enhance gene expression in mouse muscle compared to naked DNA.[50] Lower N/P ratios of 0.5 and 1 were used in the study. The intramuscular delivery of plasmid encoding antigens can induce antibody and cytotoxic T-lymphocyte (CTL) responses, which is desirable for DNA vaccination.

Lung

Gene delivery to the lung has the potential to treat several life-threatening genetic diseases such as cystic fibrosis (CF) and α_1-antitrypsin (α_1AT) deficiency. Cystic fibrosis is the result of mutations in a single gene (CF transmembrane conductance regulator, CFTR), and the gene delivery target is the airway epithelial cell. α_1AT is another lung disease resulting from a single gene mutation; however, α_1AT is a secreted product that does not need to be expressed in lung epithelium. Pulmonary gene delivery can be achieved through different administration methods, such as lung instillation, aerosol delivery, and intravenous injection. The method of application greatly influences the extent of gene expression in the lung as well as the cell type that is transfected. Instillation usually leads to significant but inhomogeneous gene expression in epithelial cells lining the bronchioles and distal airways, while aerosol delivery enables a more uniform distribution of gene expression.[70,71] By intravenous injection, transfection is mainly restricted to pulmonary endothelium.[72] Cationic polymers such as PEI and PAMAM have been proposed for pulmonary gene therapy.[73,74] Several problems including avoidance of biological barriers, suppression of immune response, and increase of transfection efficiency and specificity need to be solved before successful pulmonary gene therapy becomes a reality.

Tumor

Cancer gene therapy can take the form of the administration of a good copy of a mutated tumor suppressor gene, a gene encoding an enzyme that activates an anticancer prodrug, or a gene that encodes an antigen designed to generate a protective immune response.[75] There are two general methods for tumor treatment. One is the direct injection of DNA particles into arterial blood vessels that supply the tumor. Another method is the direct injection of the particles into the tumor.[76] In this case, the mobility of the particles within the tissue is very important to the transfection efficiency and is affected by the particle size, charge, and concentration. Liposomal and naked DNA formulations are two intensively investigated intratumoral gene delivery systems. Polyethylenimine,[77] poly(α-(4-aminobutyl)-L-glycolic acid),[78] and polyamidoamine[79] have been the polymeric gene carriers studied for cancer gene therapy.

Liver

Gene delivery to the liver involves intravenous administration of DNA containing formulations, either into a systemic or the portal vein. The stability of DNA complexes in the blood and the dilution of DNA dose by the blood are two major challenges for this application. Interestingly, naked DNA has been found to transfect hepatocytes by intravenous delivery under some special circumstances. Hydrodynamic injection of a large volume of naked DNA solution via the tail vein in 5–8 sec can induce potent gene expression in internal organs, especially the liver.[5] Significant gene expression can also be achieved in the liver by transiently restricting blood flow through the liver immediately after peripheral intravenous injection of naked DNA.[80] The hypothesis is that naked DNA is taken up by hepatocytes through receptor-mediated pathway and pressure can increase cellular uptake. However, these methods are not practical in clinical application. Condensed systems, such as poly-L-lysine[81] and PEI,[82] have been reported to achieve transfection in the liver.

Brain

Gene therapy of the brain is hindered by the presence of the blood–brain barrier (BBB), which only allows permeation of small lipophilic molecules. Exogenous genes have been expressed in the brain after invasive routes of administration, such as craniotomy[83] or intracarotid arterial infusion of noxious agents causing BBB disruption.[84] Taking advantage of receptor-mediated transport, a noninvasive gene delivery system using transferrin as receptor to cross the BBB has been described.[85] PEI has been shown to be a versatile and efficient vector in the mammalian brain.[86,87] As neurons are capable of taking up exogenous particulates from the muscles they innervate, a noninvasive strategy for gene delivery to the brain stem has been developed via peripheral intramuscular injection of PEI/DNA complexes. It is hypothesized that the complexes can reach the neuronal cell bodies in the brain stem by retrograde axonal transport after being internalized by nerve terminals in the tongue muscle.[88]

Gastrointestinal Tract

Oral gene delivery is by far the most attractive of administration. The efficiency of chitosan–DNA nanoparticles has been explored as an oral gene delivery system in a peanut allergy vaccine model.[44] The study demonstrates that transgene expression can be achieved in the small intestine of the mouse after oral administration of chitosan/DNA nanoparticles. Using pCMVArh2 plasmid encoding a dominant anaphylaxis-inducing antigen identified in mice sensitized with peanut, nanoparticle immunization elicits a relatively high IgG2a immune response and protects the animals from an allergic challenge.

CONCLUSION

Successful nonviral gene therapy will rely on the development of efficient vectors that can overcome the multitude of delivery barriers, including targeted cellular uptake, cytoplasmic release, and nuclear transport. An ideal nonviral vector would be nonimmunogenic, nontoxic, biodegradable, able to condense DNA into nanoparticles, and amenable to conjugation of ligands, endosomolytic agents, and NLS peptides. The DNA complex should possess stability before and after cellular uptake, but instability after nuclear transport. Although significant progress has been made in the development of DNA condensing agents, poor understanding of the relationship between the structure of gene carriers and transfection efficiency has rendered the design of nonviral gene delivery systems suboptimal. Systematic mechanistic studies to identify the rate-limiting steps in the nonviral gene transfer process will undoubtedly produce the next generation of efficient and safe nonviral vectors.

REFERENCES

1. Lee, R.J.; Huang, L. Folate-targeted, anionic liposome-entrapped polylysine-condensed DNA for tumor cell-specific gene transfer. J. Biol. Chem. **1996**, *271*, 8481–8487.
2. Marshall, E. Gene therapy death prompts review of adenovirus vector. Science **1999**, *286*, 2244–2245.
3. Felgner, J.H.; Kumar, R.; Sridhar, C.N.; Wheeler, C.J.; Tsai, Y.J.; Border, R.; Ramsey, P.; Martin, M.; Felgner, P.L. Enhanced gene delivery and mechanism studies with a novel series of cationic lipid formulations. J. Biol. Chem. **1994**, *269*, 2550–2561.
4. Roy, I.; Mitra, S.; Maitra, A.; Mozumdar, S. Calcium phosphate nanoparticles as novel non-viral vectors for targeted gene delivery. Int. J. Pharm. **2003**, *250*, 25–33.
5. Budker, V.; Budker, T.; Zhang, G.; Subbotin, V.; Loomis, A.; Wolff, J.A. Hypothesis: Naked plasmid DNA is taken up by cells in vivo by a receptor-mediated process. J. Gene Med. **2000**, *2*, 76–88.
6. Bloomfield, V.A. DNA condensation. Curr. Opin. Struck. Biol. **1996**, *6*, 334–341.
7. Vijayanathan, V.; Thomas, T.; Thomas, T.J. DNA nanoparticles and development of DNA delivery vehicles for gene therapy. Biochemistry **2002**, *41*, 14085–14094.
8. Fang, Y.; Hoh, J.H. Early intermediates in spermidine-induced DNA condensation on the surface of mica. J. Am. Chem. Soc. **1998**, *120*, 8903–8909.
9. Bottcher, C.; Endisch, C.; Fuhrhop, J.H.; Catterall, C.; Eaton, M. High-yield preparation of oligomeric C-type DNA toroids and their characterization by cryoelectron microscopy. J. Am. Chem. Soc. **1998**, *120*, 12–17.
10. Widom, J.; Baldwin, R.L. Monomolecular condensation of lambda-DNA induced by cobalt hexamine. Biopolymers **1983**, *22*, 1595–1620.
11. Hud, N.V.; Downing, K.H.; Balhorn, R. A constant radius of curvature model for the organization of DNA in toroidal condensates. Proc. Natl. Acad. Sci. U. S. A. **1995**, *92*, 3581–3585.
12. Rouzina, I.; Bloomfield, V.A. DNA bending by small, mobile multivalent cations. Biophys. J. **1998**, *74*, 3152–3164.
13. Golan, R.; Pietrasanta, L.I.; Hsieh, W.; Hansma, H.G. DNA toroids: Stages in condensation. Biochemistry **1999**, *38*, 14069–14076.
14. Bloomfield, V.A. DNA condensation by multivalent cations. Biopolymers **1997**, *44*, 269–282.
15. Li, A.Z.; Marx, K.A. The iso-competition point for counterion competition binding to DNA: Calculated multivalent versus monovalent cation binding equivalence. Biophys. J. **1999**, *77*, 114–122.
16. Saminathan, M.; Thomas, T.; Shirahata, A.; Pillai, C.K.; Thomas, T.J. Polyamine structural effects on the induction and stabilization of liquid crystalline DNA: Potential applications to DNA packaging, gene therapy and polyamine therapeutics. Nucleic Acids Res. **2002**, *30*, 3722–3731.
17. Hud, N.V.; Downing, K.H. Cryoelectron microscopy of lambda phage DNA condensates in vitreous ice: The fine structure of DNA toroids. Proc. Natl. Acad. Sci. U. S. A. **2001**, *98*, 14925–14930.
18. Saminathan, M.; Antony, T.; Shirahata, A.; Sigal, L.H.; Thomas, T.; Thomas, T.J. Ionic and structural specificity effects of natural and synthetic polyamines on the aggregation and resolubilization of

single-, double-, and triple-stranded DNA. Biochemistry **1999**, *38*, 3821–3830.
19. Nishikawa, M.; Huang, L. Nonviral vectors in the new millennium: Delivery barriers in gene transfer. Hum. Gene Ther. **2001**, *12*, 861–870.
20. Zabner, J.; Fasbender, A.J.; Moninger, T.; Poellinger, K.A.; Welsh, M.J. Cellular and molecular barriers to gene transfer by a cationic lipid. J. Biol. Chem. **1995**, *270*, 18997–19007.
21. Midoux, P.; Mendes, C.; Legrand, A.; Raimond, J.; Mayer, R.; Monsigny, M.; Roche, A.C. Specific gene transfer mediated by lactosylated poly-L-lysine into hepatoma cells. Nucleic Acids Res. **1993**, *21*, 871–878.
22. El Ouahabi, A.; Thiry, M.; Pector, V.; Fuks, R.; Ruysschaert, J.M.; Vandenbranden, M. The role of endosome destabilizing activity in the gene transfer process mediated by cationic lipids. FEBS Lett. **1997**, *414*, 187–192.
23. Sosnowski, B.A.; Gonzalez, A.M.; Chandler, L.A.; Buechler, Y.J.; Pierce, G.F.; Baird, A. Targeting DNA to cells with basic fibroblast growth factor (FGF2). J. Biol. Chem. **1996**, *271*, 33647–33653.
24. Boussif, O.; Lezoualc'h, F.; Zanta, M.A.; Mergny, M.D.; Scherman, D.; Demeneix, B.; Behr, J.P. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: Polyethylenimine. Proc. Natl. Acad. Sci. U. S. A. **1995**, *92*, 7297–7301.
25. Xu, Y.; Szoka, F.C., Jr. Mechanism of DNA release from cationic liposome/DNA complexes used in cell transfection. Biochemistry **1996**, *35*, 5616–5623.
26. Lechardeur, D.; Sohn, K.J.; Haardt, M.; Joshi, P.B.; Monck, M.; Graham, R.W.; Beatty, B.; Squire, J.; O'Brodovich, H.; Lukacs, G.L. Metabolic instability of plasmid DNA in the cytosol: A potential barrier to gene transfer. Gene Ther. **1999**, *6*, 482–497.
27. Branden, L.J.; Mohamed, A.J.; Smith, C.I. A peptide nucleic acid-nuclear localization signal fusion that mediates nuclear transport of DNA. Nat. Biotechnol. **1999**, *17*, 784–787.
28. Ziemienowicz, A.; Gorlich, D.; Lanka, E.; Hohn, B.; Rossi, L. Import of DNA into mammalian nuclei by proteins originating from a plant pathogenic bacterium. Proc. Natl. Acad. Sci. U. S. A. **1999**, *96*, 3729–3733.
29. Maruyama, A.; Watanabe, H.; Ferdous, A.; Katoh, M.; Ishihara, T.; Akaike, T. Characterization of interpolyelectrolyte complexes between double-stranded DNA and polylysine comb-type copolymers having hydrophilic side chains. Bioconjug. Chem. **1998**, *9*, 292–299.
30. Choi, Y.H.; Liu, F.; Kim, J.S.; Choi, Y.K.; Park, J.S.; Kim, S.W. Polyethylene glycol-grafted poly-L-lysine as polymeric gene carrier. J. Control. Release **1998**, *54*, 39–48.
31. Wolfert, M.A.; Schacht, E.H.; Toncheva, V.; Ulbrich, K.; Nazarova, O.; Seymour, L.W. Characterization of vectors for gene therapy formed by self-assembly of DNA with synthetic block copolymers. Hum. Gene Ther. **1996**, *7*, 2123–2133.
32. Jeong, J.H.; Park, T.G. Poly(L-lysine)-g-poly(D-L-lactic-co-glycolic acid) micelles for low cytotoxic biodegradable gene delivery carriers. J. Control. Release **2002**, *82*, 159–166.
33. Wightman, L.; Kircheis, R.; Rossler, V.; Carotta, S.; Ruzicka, R.; Kursa, M.; Wagner, E. Different behavior of branched and linear polyethylenimine for gene delivery in vitro and in vivo. J. Gene Med. **2001**, *3*, 362–372.
34. Nguyen, H.K.; Lemieux, P.; Vinogradov, S.V.; Gebhart, C.L.; Guerin, N.; Paradis, G.; Bronich, T.K.; Alakhov, V.Y.; Kabanov, A.V. Evaluation of polyether-polyethyleneimine graft copolymers as gene transfer agents. Gene Ther. **2000**, *7*, 126–138.
35. Gosselin, M.A.; Guo, W.; Lee, R.J. Efficient gene transfer using reversibly cross-linked low molecular weight polyethylenimine. Bioconjug. Chem. **2001**, *12*, 989–994.
36. Kukowska-Latallo, J.F.; Bielinska, A.U.; Johnson, J.; Spindler, R.; Tomalia, D.A.; Baker, J.R., Jr. Efficient transfer of genetic material into mammalian cells using Starburst polyamidoamine dendrimers. Proc. Natl. Acad. Sci. U. S. A. **1996**, *93*, 4897–4902.
37. Zhang, Z.Y.; Smith, B.D. High-generation polycationic dendrimers are unusually effective at disrupting anionic vesicles: Membrane bending model. Bioconjug. Chem. **2000**, *11*, 805–814.
38. Lim, Y.B.; Han, S.O.; Kong, H.U.; Lee, Y.; Park, J.S.; Jeong, B.; Kim, S.W. Biodegradable polyester, poly[alpha-(4-aminobutyl)-L-glycolic acid], as a non-toxic gene carrier. Pharm. Res. **2000**, *17*, 811–816.
39. Maheshwari, A.; Mahato, R.I.; McGregor, J.; Han, S.; Samlowski, W.E.; Park, J.S.; Kim, S.W. Soluble biodegradable polymer-based cytokine gene delivery for cancer treatment. Mol. Ther. **2000**, *2*, 121–130.
40. Cherng, J.Y.; van de Wetering, P.; Talsma, H.; Crommelin, D.J.; Hennink, W.E. Effect of size and serum proteins on transfection efficiency of poly ((2-dimethylamino)ethyl methacrylate)-plasmid nanoparticles. Pharm. Res. **1996**, *13*, 1038–1042.
41. vandeWetering, P.; Cherng, J.Y.; Talsma, H.; Hennink, W.E. Relation between transfection efficiency and cytotoxicity of poly(2-(dimethylami-

no)ethyl methacrylate)/plasmid complexes. J. Control. Release **1997**, *49*, 59–69.
42. van de Wetering, P.; Moret, E.E.; Schuurmans-Nieuwenbroek, N.M.; van Steenbergen, M.J.; Hennink, W.E. Structure-activity relationships of water-soluble cationic methacrylate/methacrylamide polymers for nonviral gene delivery. Bioconjug. Chem. **1999**, *10*, 589–597.
43. MacLaughlin, F.C.; Mumper, R.J.; Wang, J.J.; Tagliaferri, J.M.; Gill, I.; Hinchcliffe, M.; Rolland, A.P. Chitosan and depolymerized chitosan oligomers as condensing carriers for in vivo plasmid delivery. J. Control. Release **1998**, *56*, 259–272.
44. Roy, K.; Mao, H.Q.; Huang, S.K.; Leong, K.W. Oral gene delivery with chitosan–DNA nanoparticles generates immunologic protection in a murine model of peanut allergy. Nat. Med. **1999**, *5*, 387–391.
45. Richardson, S.C.; Kolbe, H.V.; Duncan, R. Potential of low molecular mass chitosan as a DNA delivery system: Biocompatibility, body distribution and ability to complex and protect DNA. Int. J. Pharm. **1999**, *178*, 231–243.
46. Thanou, M.; Florea, B.I.; Geldof, M.; Junginger, H.E.; Borchard, G. Quaternized chitosan oligomers as novel gene delivery vectors in epithelial cell lines. Biomaterials **2002**, *23*, 153–159.
47. Lee, K.Y.; Kwon, I.C.; Kim, Y.H.; Jo, W.H.; Jeong, S.Y. Preparation of chitosan self-aggregates as a gene delivery system. J. Control. Release **1998**, *51*, 213–220.
48. Kiang, T.; Wen, J.; Lim, H.; Leong, K.W. Degree of Deacetylation of Chitosan: Effect on Gene Transfection Efficiency In vivo. Proceedings of 29th Annual Meeting of the Controlled Release Society, Seoul, Korea, July 20–25, 2002.
49. Wang, J.; Mao, H.Q.; Leong, K.W. A novel biodegradable gene carrier based on polyphosphoester. J. Am. Chem. Soc. **2001**, *123*, 9480–9481.
50. Wang, J.; Zhang, P.C.; Mao, H.Q.; Leong, K.W. Enhanced gene expression in mouse muscle by sustained release of plasmid DNA using PPE-EA as a carrier. Gene Ther. **2002**, *9*, 1254–1261.
51. Wen, J.; Mao, H.Q.; Lin, K.; Li, W.; Leong, K.W. Novel Biodegradable Polyphosphoester Micelles as Gene Carriers. Proceedings of 3rd Annual Meeting of the American Society of Gene Therapy, Denver, CO, May 31–June 4, 2000.
52. Davis, S.S. Biomedical applications of nanotechnology—Implications for drug targeting and gene therapy. Trends Biotechnol. **1997**, *15*, 217–224.
53. Poznansky, M.J.; Juliano, R.L. Biological approaches to the controlled delivery of drugs: A critical review. Pharmacol. Rev. **1984**, *36*, 277–336.
54. Patel, H.M. Serum opsonins and liposomes: Their interaction and opsonophagocytosis. Crit. Rev. Ther. Drug Carr. Syst. **1992**, *9*, 39–90.
55. Wagner, E.; Cotten, M.; Foisner, R.; Birnstiel, M.L. Transferrin-polycation-DNA complexes: The effect of polycations on the structure of the complex and DNA delivery to cells. Proc. Natl. Acad. Sci. U. S. A. **1991**, *88*, 4255–4259.
56. Puls, R.; Minchin, R. Gene transfer and expression of a non-viral polycation-based vector in $CD4^+$ cells. Gene Ther. **1999**, *6*, 1774–1778.
57. Diebold, S.S.; Lehrmann, H.; Kursa, M.; Wagner, E.; Cotten, M.; Zenke, M. Efficient gene delivery into human dendritic cells by adenovirus polyethylenimine and mannose polyethylenimine transfection. Hum. Gene Ther. **1999**, *10*, 775–786.
58. Mahato, R.I.; Takemura, S.; Akamatsu, K.; Nishikawa, M.; Takakura, Y.; Hashida, M. Physicochemical and disposition characteristics of antisense oligonucleotides complexed with glycosylated poly(L-lysine). Biochem. Pharmacol. **1997**, *53*, 887–895.
59. Kollen, W.J.; Mulberg, A.E.; Wei, X.; Sugita, M.; Raghuram, V.; Wang, J.; Foskett, J.K.; Glick, M.C.; Scanlin, T.F. High-efficiency transfer of cystic fibrosis transmembrane conductance regulator cDNA into cystic fibrosis airway cells in culture using lactosylated polylysine as a vector. Hum. Gene Ther. **1999**, *10*, 615–622.
60. Leamon, C.P.; Weigl, D.; Hendren, R.W. Folate copolymer-mediated transfection of cultured cells. Bioconjug. Chem. **1999**, *10*, 947–957.
61. Kim, J.S.; Maruyama, A.; Akaike, T.; Kim, S.W. Terplex DNA delivery system as a gene carrier. Pharm. Res. **1998**, *15*, 116–121.
62. Harbottle, R.P.; Cooper, R.G.; Hart, S.L.; Ladhoff, A.; McKay, T.; Knight, A.M.; Wagner, E.; Miller, A.D.; Coutelle, C. An RGD-oligolysine peptide: A prototype construct for integrin-mediated gene delivery. Hum. Gene Ther. **1998**, *9*, 1037–1047.
63. Ludtke, J.J.; Zhang, G.; Sebestyen, M.G.; Wolff, J.A. A nuclear localization signal can enhance both the nuclear transport and expression of 1 kb DNA. J. Cell. Sci. **1999**, *112* (Pt. 12), 2033–2041.
64. Chan, C.K.; Jans, D.A. Enhancement of polylysine-mediated transferrinfection by nuclear localization sequences: Polylysine does not function as a nuclear localization sequence. Hum. Gene Ther. **1999**, *10*, 1695–1702.
65. van Dijk-Wolthuis, W.N.; van de Wetering, P.; Hinrichs, W.L.; Hofmeyer, L.J.; Liskamp, R.M.; Crommelin, D.J.; Hennink, W.E. A versatile method for the conjugation of proteins and peptides to poly[2-(dimethylamino)ethyl methacrylate]. Bioconjug. Chem. **1999**, *10*, 687–692.
66. Sebestyen, M.G.; Ludtke, J.J.; Bassik, M.C.; Zhang,

G.; Budker, V.; Lukhtanov, E.A.; Hagstrom, J.E.; Wolff, J.A. DNA vector chemistry: The covalent attachment of signal peptides to plasmid DNA. Nat. Biotechnol. **1998**, *16*, 80–85.

67. Zanta, M.A.; Belguise-Valladier, P.; Behr, J.P. Gene delivery: A single nuclear localization signal peptide is sufficient to carry DNA to the cell nucleus. Proc. Natl. Acad. Sci. U. S. A. **1999**, *96*, 91–96.
68. Wolff, J.A.; Malone, R.W.; Williams, P.; Chong, W.; Acsadi, G.; Jani, A.; Felgner, P.L. Direct gene transfer into mouse muscle in vivo. Science **1990**, *247*, 1465–1468.
69. Mumper, R.J.; Wang, J.; Klakamp, S.L.; Nitta, H.; Anwer, K.; Tagliaferri, F.; Rolland, A.P. Protective interactive noncondensing (PINC) polymers for enhanced plasmid distribution and expression in rat skeletal muscle. J. Control. Release **1998**, *52*, 191–203.
70. Uyechi, L.S.; Gagne, L.; Thurston, G.; Szoka, F.C., Jr. Mechanism of lipoplex gene delivery in mouse lung: Binding and internalization of fluorescent lipid and DNA components. Gene Ther. **2001**, *8*, 828–836.
71. Guillaume, C.; Delepine, P.; Droal, C.; Montier, T.; Tymen, G.; Claude, F. Aerosolization of cationic lipid-DNA complexes: Lipoplex characterization and optimization of aerosol delivery conditions. Biochem. Biophys. Res. Commun. **2001**, *286*, 464–471.
72. Bragonzi, A.; Boletta, A.; Biffi, A.; Muggia, A.; Sersale, G.; Cheng, S.H.; Bordignon, C.; Assael, B.M.; Conese, M. Comparison between cationic polymers and lipids in mediating systemic gene delivery to the lungs. Gene Ther. **1999**, *6*, 1995–2004.
73. Trubetskoy, V.S.; Wong, S.C.; Subbotin, V.; Budker, V.G.; Loomis, A.; Hagstrom, J.E.; Wolff, J.A. Recharging cationic DNA complexes with highly charged polyanions for in vitro and in vivo gene delivery. Gene Ther. **2003**, *10*, 261–271.
74. Kukowska-Latallo, J.F.; Raczka, E.; Quintana, A.; Chen, C.; Rymaszewski, M.; Baker, J.R., Jr. Intravascular and endobronchial DNA delivery to murine lung tissue using a novel, nonviral vector. Hum. Gene Ther. **2000**, *11*, 1385–1395.
75. Brown, M.D.; Schatzlein, A.G.; Uchegbu, I.F. Gene delivery with synthetic (non viral) carriers. Int. J. Pharm. **2001**, *229*, 1–21.
76. Merdan, T.; Kopecek, J.; Kissel, T. Prospects for cationic polymers in gene and oligonucleotide therapy against cancer. Adv. Drug Deliv. Rev. **2002**, *54*, 715–758.
77. Aoki, K.; Furuhata, S.; Hatanaka, K.; Maeda, M.; Remy, J.S.; Behr, J.P.; Terada, M.; Yoshida, T. Polyethylenimine-mediated gene transfer into pancreatic tumor dissemination in the murine peritoneal cavity. Gene Ther. **2001**, *8*, 508–514.
78. Maheshwari, A.; Han, S.; Mahato, R.I.; Kim, S.W. Biodegradable polymer-based interleukin-12 gene delivery: Role of induced cytokines, tumor infiltrating cells and nitric oxide in anti-tumor activity. Gene Ther. **2002**, *9*, 1075–1084.
79. Sato, N.; Kobayashi, H.; Saga, T.; Nakamoto, Y.; Ishimori, T.; Togashi, K.; Fujibayashi, Y.; Konishi, J.; Brechbiel, M.W. Tumor targeting and imaging of intraperitoneal tumors by use of antisense oligo-DNA complexed with dendrimers and/or avidin in mice. Clin. Cancer Res. **2001**, *7*, 3606–3612.
80. Liu, F.; Huang, L. Improving plasmid DNA-mediated liver gene transfer by prolonging its retention in the hepatic vasculature. J. Gene Med. **2001**, *3*, 569–576.
81. Zhang, X.; Collins, L.; Sawyer, G.J.; Dong, X.; Qiu, Y.; Fabre, J.W. In vivo gene delivery via portal vein and bile duct to individual lobes of the rat liver using a polylysine-based nonviral DNA vector in combination with chloroquine. Hum. Gene Ther. **2001**, *12*, 2179–2190.
82. Yamazaki, Y.; Nango, M.; Matsuura, M.; Hasegawa, Y.; Hasegawa, M.; Oku, N. Polycation liposomes, a novel nonviral gene transfer system, constructed from cetylated polyethylenimine. Gene Ther. **2000**, *7*, 1148–1155.
83. Culver, K.W.; Ram, Z.; Wallbridge, S.; Ishii, H.; Oldfield, E.H.; Blaese, R.M. In vivo gene transfer with retroviral vector-producer cells for treatment of experimental brain tumors. Science **1992**, *256*, 1550–1552.
84. Nilaver, G.; Muldoon, L.L.; Kroll, R.A.; Pagel, M.A.; Breakefield, X.O.; Davidson, B.L.; Neuwelt, E.A. Delivery of herpesvirus and adenovirus to nude rat intracerebral tumors after osmotic blood–brain barrier disruption. Proc. Natl. Acad. Sci. U. S. A. **1995**, *92*, 9829–9833.
85. Shi, N.; Pardridge, W.M. Noninvasive gene targeting to the brain. Proc. Natl. Acad. Sci. U. S. A. **2000**, *97*, 7567–7572.
86. Abdallah, B.; Hassan, A.; Benoist, C.; Goula, D.; Behr, J.P.; Demeneix, B.A. A powerful nonviral vector for in vivo gene transfer into the adult mammalian brain: Polyethylenimine. Hum. Gene Ther. **1996**, *7*, 1947–1954.
87. Goula, D.; Remy, J.S.; Erbacher, P.; Wasowicz, M.; Levi, G.; Abdallah, B.; Demeneix, B.A. Size, diffusibility and transfection performance of linear PEI/DNA complexes in the mouse central nervous system. Gene Ther. **1998**, *5*, 712–717.
88. Wang, S.; Ma, N.; Gao, S.J.; Yu, H.; Leong, K.W. Transgene expression in the brain stem effected by intramuscular injection of polyethylenimine/DNA complexes. Molec. Ther. **2001**, *3*, 658–664.

Polymer Nanowires Conjugated by Controlled Chain Polymerization

Yuji Okawa
Masakazu Aono
National Institute for Materials Science, Ibaraki, Japan

INTRODUCTION

The present-day silicon-based electronic devices have grown more powerful as their basic subunit, the transistor, has shrunk in size. However, the laws of quantum mechanics and the limitations of fabrication techniques may soon prevent further reduction. Thus we must develop a novel device concept, which can be applied at the nanometer scale. This exploration of a novel device concept beyond silicon-based transistors is analogous to the search for a replacement for the vacuum tube half a century ago, and its discovery will make a profound impact on society in this century. Many ideas have been proposed for the novel device concept, and one promising possibility is the construction of novel electronic devices using organic molecules.[1–14] However, even if we could make individual molecules that function as single-molecule transistors or diodes, the next problem would be the absence of good methods by which to arrange them in the desired pattern and to interconnect them with one another. Regarding this problem, it is necessary to develop a method of fabricating conductive wires of nanometer width (nanowires) at designated positions.

We usually use metals for wires whose width is larger than the order of micrometers. However, in the case of nanowires, the use of metals becomes much more difficult because of their poor stability and difficulties in their fabrication. If we want to use metals for nanowires, we must devise some method of stabilizing them.[15] Hence many materials other than metals have been considered as candidates for nanowires, including inorganic compounds,[16] carbon nanotubes,[4,8,9,12,14,17–19] DNA,[20–24] and other conductive organic molecules or polymers,[25–30] although we must conduct further basic studies before we can put them to practical use.

In this article, we will present a method of fabricating a conjugated polymer nanowire at designated positions by controlling linear chain polymerization using the probe tip of a scanning tunneling microscope (STM).[29,30] A demonstration is presented for a self-ordered monomolecular layer of 10,12-pentacosadiynoic acid and 10,12-nonacosadiynoic acid, which are amphiphilic diacetylene compounds, adsorbed on a graphite substrate. We have succeeded in controlling the initiation and termination of chain polymerization with a spatial precision on the order of 1 nm. The obtained polydiacetylene is a conjugated polymer, so that we can expect that it will become conductive when excess electrons or holes are supplied to it from the surroundings.

SELF-ORDERED MOLECULAR LAYER

The monomer molecules we use are diacetylene compounds with the general formula R-C≡C—C≡C—R′, where C≡C—C≡C is the diacetylene moiety, and R and R′ are substituent groups. It is well known that this molecular species in the form of solid crystals[31] or Langmuir–Brodgett films[32] polymerizes into polydiacetylene compounds on appropriate stimulation such as heating or ultraviolet irradiation (Fig. 1). Several studies have dealt with STM observations of physisorbed layers of various diacetylene compounds on graphite surfaces.[33–35] According to those studies, a variety of molecular arrangements are formed depending on the species of substituent groups, the method of film preparation, and the incorporation of foreign molecules.

In our work, we chose 10,12-pentacosadiynoic acid [$CH_3(CH_2)_{11}$—C≡C—C≡C—$(CH_2)_8COOH$] and 10,12-nonacosadiynoic acid [$CH_3(CH_2)_{15}$—C≡C—C≡C—$(CH_2)_8COOH$] as monomer molecules. These two species gave essentially the same results. To prepare their thin films, the monomer molecules were dissolved in chloroform and the solution was applied onto the surface of purified water. After the evaporation of chloroform, the thin films of monomer molecules on the water surface were transferred to a freshly cleaved surface of highly oriented pyrolytic graphite (HOPG) by nearly horizontal dipping.

The STM images were recorded using the Digital Instruments NanoScope STM system under ambient conditions in the constant current mode. A typical STM image of the 10,12-pentacosadiynoic acid layer is shown in Fig. 2a. The image consists of parallel bright lines

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009444

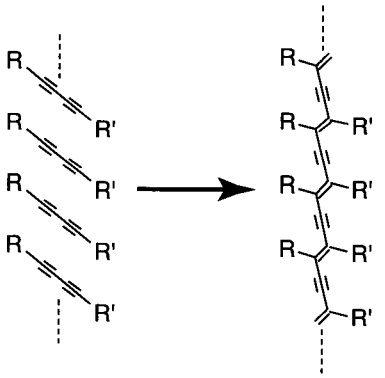


Fig. 1 Polymerization of diacetylene compounds into polydiacetylene compounds.

separated by two different alternating spacings of about 3.0 and 3.8 nm, indicating that the 10,12-pentacosadiynoic acid molecules on the graphite surface are self-ordered after the transfer from the water surface. A magnified STM image of the layer is shown in Fig. 2b, in which individual molecules are resolved. Based on such STM images, we proposed the molecular arrangement shown in Fig. 2c. The molecules are aligned in phase to form straight chains and the chains are arranged such that the COOH end groups of a chain are opposite to those of a neighboring chain. Each bright line in Fig. 2a corresponds to the linear array of diacetylene moieties in Fig. 2c because the multiple bonds are, in general, observed to be higher than single bonds in STM images.[36]

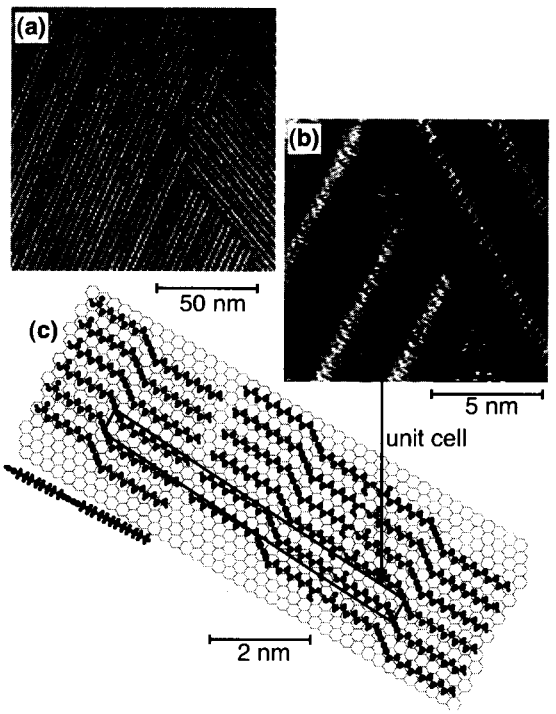


Fig. 2 (a) Typical STM image of 10,12-pentacosadiynoic acid layer on a graphite surface (sample bias voltage $V_S = -1.0$ V; tunneling current $I_t = 0.07$ nA). (b) Magnified STM image of the layer ($V_S = +0.5$ V; $I_t = 1.0$ nA). The arrows in the image indicate the main crystal axes of graphite. (c) Top and side views of the proposed model for molecular arrangement. (From Ref. [30]; © American Institute of Physics, 2001.) *(View this art in color at www.dekker.com.)*

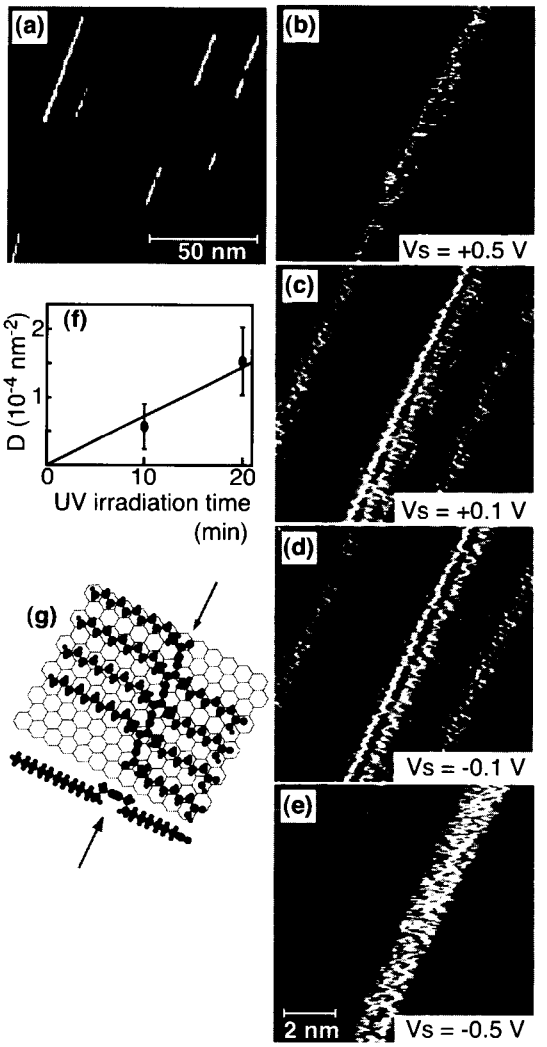


Fig. 3 (a) An STM image obtained after ultraviolet irradiation ($V_S = -1.0$ V; $I_t = 0.07$ nA). (b–e) Magnified STM images of a polymer obtained at the same tunneling current ($I_t = 0.1$ nA) and various sample biases V_S. (f) Plots of the number density of the polymer D against the duration of ultraviolet irradiation. (g) The proposed structural model of a polymer. The arrows indicate the backbone of polydiacetylene. (From Ref. [30]; © American Institute of Physics, 2001.) *(View this art in color at www.dekker.com.)*

At appropriately small sample biases and high tunneling currents, the STM image of the graphite substrate can be observed. On comparing such an image with the images of the molecular layer taken previously, it was found that the alkyl side chains of each molecule are oriented parallel to the main crystal axis of graphite. For many other molecules on graphite surfaces, it has been reported that their alkyl chains are also oriented parallel to the main crystal axis of graphite.[34,36–40] In Fig. 2b, a modulation of the image contrast is observed as a moiré pattern because of the lattice mismatch between the molecular layer and the graphite substrate. The periodicity of the moiré modulation indicates that the spacing between neighboring molecules in the molecular layer is 0.47 nm.

PHOTOPOLYMERIZATION

First, we examined whether such chain polymerization could be induced by ultraviolet irradiation in the monomer film described above, and how the created polymers are observed in STM images. For this purpose, the molecular layer was irradiated with ultraviolet light from a low-pressure mercury lamp (254 nm wavelength and 1.3 mW/cm^2 power density at the sample position) for a period of 20 min.

Fig. 3a shows a typical STM image of the 10,12-pentacosadiynoic acid layer after ultraviolet irradiation. As we can see, very bright lines appeared. One of these very bright lines is imaged on a magnified scale at various sample biases, as shown in Fig. 3b–e. The number density of the very bright lines increased almost linearly with increasing duration of ultraviolet irradiation, as shown in Fig. 3f. Hence these very bright lines are considered to be polydiacetylenes, created by the photopolymerization of the linear array of diacetylene moieties.

The moiré modulation described above was also observed for the polymer, as seen in Fig. 3c. An important point is that the moiré pattern in the created polymer is the same as that in the monomer array. This means that the spacing between the alkyl side chains in the polymer is the same as that in the monomer—0.47 nm. This also indicates that the orientation of the alkyl chains is maintained parallel to the main crystal axis of graphite even after polymerization. Based on these results, we proposed the structural model for the polymer, as shown in Fig. 3g. In this model, the polydiacetylene backbone is raised so that the alkyl side chains are parallel to the main crystal axis of the graphite substrate.

The most plausible mechanism of photopolymerization is illustrated in Fig. 4, by analogy with the case of bulk crystals of diacetylene compounds.[41] In the array of

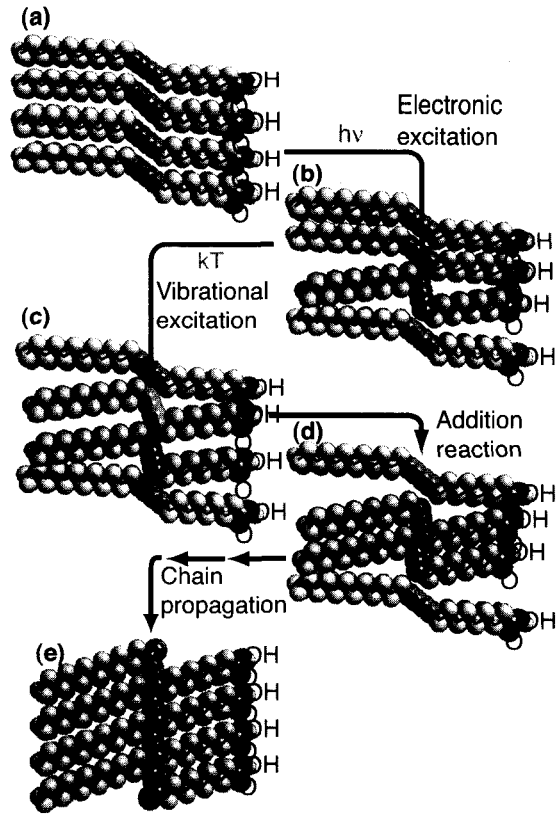


Fig. 4 Illustration of the mechanism of photopolymerization. (a) Array of the monomer molecules. (b) Diradical formation by photoabsorption. (c) Approach of a neighboring diacetylene moiety by vibrational excitation. (d) Dimer formation by an addition reaction. (e) Extended chain propagation. (From Ref. [30]; © American Institute of Physics, 2001.) (*View this art in color at www.dekker.com.*)

monomer molecules (Fig. 4a), the diacetylene moiety of one of the molecules is excited by photoabsorption into an excited state. As shown in Fig. 4b, a diradical state in the *trans* form with an unpaired electron at either end is a reasonable representation of the lowest excited $\pi\pi^*$ singlet or triplet state located at 3.7 and 3.1 eV above the $\pi\pi$ ground state, respectively.[42,43] Within the lifetime of the diradical state, if a neighboring diacetylene moiety on either side approaches the diradical because of thermal vibration (Fig. 4c), an addition reaction occurs, forming a dimer of diacetylene (Fig. 4d). Because the created dimer of diacetylene still has radicals at both ends, a similar addition reaction occurs on both ends. Hence the repetition of such addition reaction occurs and results in extended chain polymerization, as shown in Fig. 4e. Therefore we can expect that if the excited diradical state in the first step (Fig. 4b) can be created by stimulation with the STM tip instead of photoabsorption, the extended chain polymerization will follow.

CHAIN POLYMERIZATION INITIATED BY A SCANNING TUNNELING MICROSCOPE TIP

Fig. 5a shows an STM image of the 10,12-pentacosadiynoic acid layer. The same area was imaged again in Fig. 5b, in which the STM tip is scanned rapidly in the horizontal direction and slowly from the bottom to the top at a sample bias of -1 V. When the tip passed the point indicated by an arrow, a pulsed sample bias (-4 V in height and 5 μsec in width) was applied, as shown on the left-hand side of the figure. As can be seen, a very bright line appeared starting from the point where the pulsed sample bias was applied. This very bright line is identical in structure to those observed in Fig. 3. Namely, stimulation with the STM tip induced chain polymerization and created a polymer. The same area was imaged once again and the result is shown in Fig. 5c. We can see that chain polymerization occurred on both sides of the point of stimulation.

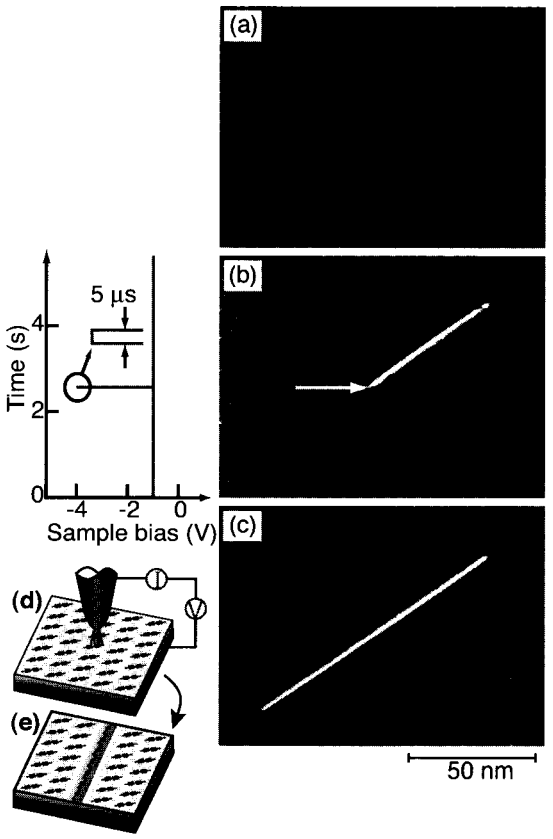


Fig. 5 (a) An STM image of the original monomolecular layer of 10,12-pentacosadiynoic acid ($V_S=-1.0$ V; $I_t=0.07$ nA). (b) Image of the same area observed from the bottom to the top, but with pulsed sample bias voltage applied at the point indicated by the arrow. (c) The same area imaged once again. (d, e) Diagrams illustrating the initiation of chain polymerization using an STM tip. (From Ref. [30]; © American Institute of Physics, 2001.) (*View this art in color at www.dekker.com.*)

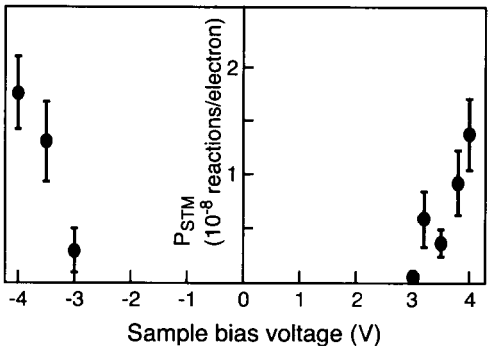


Fig. 6 Plots of the frequency of occurrence of chain polymerization P_{STM} in units of reactions per electron against the height of the pulsed sample bias voltage V_S. (From Ref. [30]; © American Institute of Physics, 2001.)

In the case of Fig. 5b, a negative pulsed sample bias (-4.0 V) was applied, and we found that chain polymerization occurred at both polarities of the pulsed sample bias. This is seen in Fig. 6, where the frequency of the occurrence of chain polymerization P_{STM} is plotted against the height of the pulsed bias voltage V_S at both polarities; these measurements were carried out using a 10,12-pentacosadiynoic acid layer at a fixed pulse width of 10 μsec while monitoring the tunneling current. As we can see, the reaction probability was symmetric with respect to the polarity of the pulsed sample bias, and the threshold voltage for polymerization was 2.9 ± 0.2 V. This value is close to the energy separation between the ππ ground state and the lowest excited ππ* triplet state of the diacetylene moiety—3.1 eV.[43] This suggests that the diacetylene moiety was excited by electrons tunneling between the tip and the substrate, which passed inelastically through the intermediate molecular layer.

As seen in Fig. 5c, the polymers created were usually terminated at domain boundaries. In other words, the chain polymerization initiated by the STM tip propagated in the domain without any additional energy other than thermal energy, and terminated when it encountered the domain boundary that acted as a kind of structural defect. Exploiting the fact that the chain reaction is terminated by a structural defect, we can also control the termination of chain polymerization by forming an artificial structural defect in advance. An example of a 10,12-nonacosadiynoic acid layer is presented in Fig. 7. First, the original monomolecular layer was imaged (Fig. 7a). Then we created an artificial defect at the center of the image, as shown in Fig. 7b, using a previously reported method.[44]

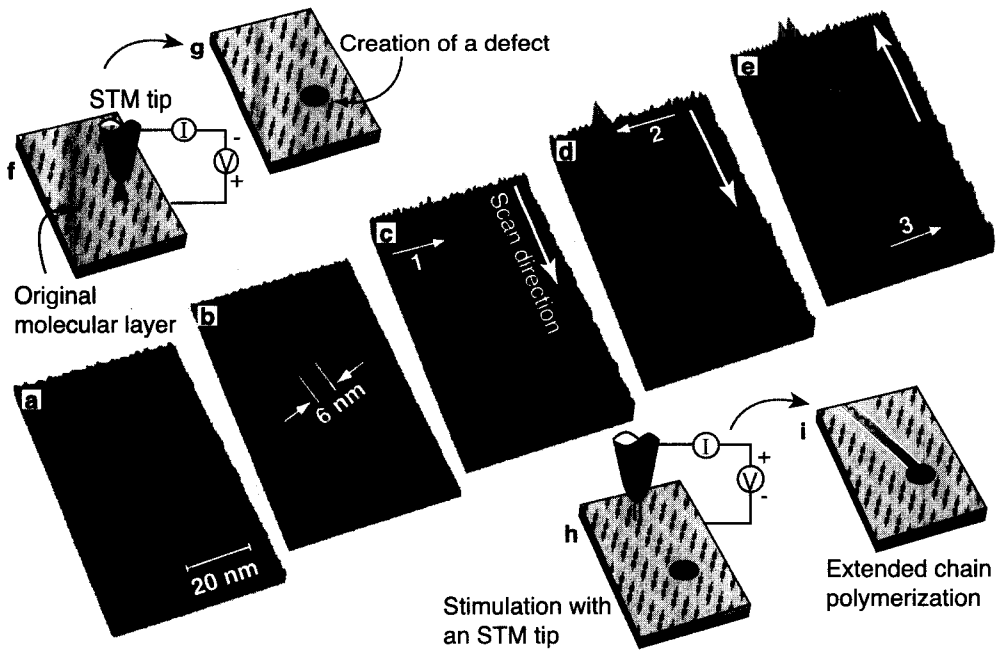


Fig. 7 (a) An STM image of the original monomolecular layer of 10,12-nonacosadiynoic acid ($V_S = -1.0$ V; $I_t = 0.07$ nA). (b) Creation of an artificial defect in advance with an STM tip. (c) First chain polymerization, initiated at the point indicated by arrow 1 and terminated at the artificial defect. (d, e) Second and third chain polymerizations initiated at the points indicated by arrows 2 and 3, respectively. (f, g) Diagrams illustrating the creation of an artificial defect in advance. (h, i) Diagrams illustrating the initiation and termination of chain polymerization. (From Ref. [29]; © Macmillan Magazines Ltd., 2001.) (*View this art in color at www.dekker.com.*)

Namely, an STM tip was placed at that position and a high pulsed sample bias (+5 V in height and 10 μsec in width) was applied (Fig. 7f and g). The created defect was a hole with a diameter of 6 nm in the graphite substrate. The same area was imaged again from the top to the bottom, as shown in Fig. 7c, but a pulsed sample bias (−4 V in height and 5 μsec in width) was applied when the scanning tip passed the point marked by arrow 1. A very bright line appeared between the point of stimulation and the artificial defect, indicating that chain polymerization was initiated at the stimulation point and terminated at the artificial defect (Fig. 7h and i). We imaged the same area once more from the top to the bottom, as shown in Fig. 7d, applying another pulsed sample bias when the tip passed the point marked by arrow 2. In this case as well, chain polymerization was initiated at the indicated point and terminated at the artificial defect. Note the chain polymerization initiated at the point marked by arrow 1 propagated on both sides of the point, as described above. In a similar manner, we created a third polymer nanowire by initiating chain polymerization at the point marked by arrow 3 in Fig. 7e.

This result shows that we can control the initiation and termination of chain polymerization with spatial precision on the order of 1 nm. Furthermore, this result also indicates that at least three polymer nanowires can be connected to an object as small as 6 nm. This means that the source, drain, and gate electrode of a nanoscale single-molecule transistor, for instance, will be connectable.

ADVANTAGES OF THE METHOD

It is well known that we can manipulate individual atoms or molecules with an STM tip,[45–48] so we should be able to create nanowires using such atom-by-atom crafting. It is also known that if a thin molecular film is stimulated by a biased STM tip, a local chemical reaction or reorientation is sometimes induced,[49–53] so that continuous modification will be possible by scanning of a biased STM tip. Although such work is essential to fundamental advances, the chain polymerization technique used in our work has several advantages.

First, the nanowires obtained are guaranteed to have perfect structures without any defects. If defects exist, the chain reaction itself will be terminated, as has been discussed above. Hence if the chain reaction proceeds, the nanowire reaction product is guaranteed to have perfect structures. For a chain reaction proceeding in a large domain without any defects, polymers longer than 500 nm have been observed. Furthermore, a single conducting polymer chain in a free state, such as a solution, generally

has a coiled conformation, which prevents the occurrence of π-electron conjugation along the whole chain. In contrast, the polydiacetylene nanowires on a graphite surface obtained in our method are straight, so they can be considered ideal π-conjugated systems. Second, the nanowires are stable at room temperature. Third, we can create a long nanowire by applying only one pulsed bias voltage at one point; the propagation speed of the chain reaction is high (we roughly estimated the propagation speed of the chain polymerization to be 0.1–1 m/sec), so that the entire process is accomplished in a short time. Fourth, the process required no additional energy for the chain propagation reaction, other than the first stimulation.

The controlled chain reaction methods have many advantages for the fabrication of nanosized structures. However, despite these advantages, reports on controlled chain reaction are few (Lopinski et al.[54] have reported another type of chain reaction—a spontaneous growth of a molecular stylene line from a silicon dangling bond).

Finally, we would like to discuss whether or not the nanowires are conductive. According to electronic energy band calculations for various polydiacetylene compounds,[55–57] their polydiacetylene backbones have almost the same band structure and are nearly independent of substituent groups. Namely, the conjugated π-electrons of the polydiacetylene backbone form valence and conduction bands, which are separated by an energy gap of about 1.6 eV.[57] Because of the existence of this band gap, the polydiacetylene backbone is expected to be a nonconductor. However, because the highest valence band and the lowest conduction band have fairly strong dispersions,[57] the polydiacetylene backbone should act as an electronically conductive nanowire on the transfer of charge from the surroundings. For example, excess electrons, created by light illumination on a polydiacetylene chain, move in response to an electronic field with an ultrahigh drift mobility,[58] and the conductivity of polydiacetylene films increases on doping with iodine.[59] The polymer nanowire on the graphite surface created by our method is also expected to be conductive if a charge transfer between the polymer and the graphite surface exists. We are now planning to measure the conductivity of the polymer nanowire.

CONCLUSION

We have demonstrated that we can initiate the linear propagation of the chain polymerization of unsaturated organic molecules at any predetermined point and terminate it at another predetermined point with a spatial precision on the order of 1 nm, using the probe tip of an STM. In our demonstration, we used a self-ordered monomolecular layer of the amphiphilic diacetylene compounds, 10,12-pentacosadiynoic acid and 10,12-nonacosadiynoic acid. Because the polydiacetylene backbone has a conjugated π-electron system, we can expect that it will function as an electrically conductive nanowire when charge is transferred from the surroundings.

In the future, this method will be useful for fabricating novel molecular devices comprising conjugated polymer nanowires, and for interconnecting novel nanoelectronic devices. Furthermore, because the obtained polymer is considered to be a long, linear, conductive nanowire with high structural perfection, it will be an ideal system for the basic study of the physics of one-dimensional conductors. We also hope that this work will provide a new experimental method for chemists investigating the kinetics and mechanisms of chain reactions.

REFERENCES

1. Aviram, A.; Ratner, M.A. Molecular rectifiers. Chem. Phys. Lett. **1974**, *29* (2), 277–283.
2. Aviram, A. Molecules for memory, logic, and amplification. J. Am. Chem. Soc. **1988**, *110* (17), 5687–5692.
3. Joachim, C.; Gimzewski, J.K. An electromechanical amplifier using a single molecule. Chem. Phys. Lett. **1997**, *265*, 353–357.
4. Tans, S.J.; Verschueren, A.R.M.; Dekker, C. Room-temperature transistor based on a single carbon nanotube. Nature **1998**, *393*, 49–52.
5. Collier, C.P.; Wong, E.W.; Belohradsky, M.; Raymo, F.M.; Stoddart, J.F.; Kuekes, P.J.; Williams, R.S.; Heath, J.R. Electronically configurable molecular-based logic gates. Science **1999**, *285*, 391–394.
6. Chen, J.; Reed, M.A.; Rawlett, A.M.; Tour, J.M. Large on–off ratios and negative differential resistance in a molecular electronic device. Science **1999**, *286*, 1550–1552.
7. Joachim, C.; Gimzewski, J.K.; Aviram, A. Electronics using hybrid-molecular and mono-molecular devices. Nature **2000**, *408*, 541–548.
8. Batchtold, A.; Hadley, P.; Nakanishi, T.; Dekker, C. Logic circuits with carbon nanotube transistors. Science **2001**, *294*, 1317–1320.
9. Liu, X.; Lee, C.; Zhou, C.; Han, J. Carbon nanotube field-effect inverters. Appl. Phys. Lett. **2001**, *79* (20), 3329–3331.
10. Xue, J.; Forrest, S.R. Organic thin-film transistors based on bis(1,2,5-thiadiazolo)-*p*-quinobis(1,3-dithiole). Appl. Phys. Lett. **2001**, *79* (22), 3714–3716.
11. Collier, C.P.; Jeppesen, J.O.; Luo, Y.; Perkins, J.; Wong, E.W.; Heath, J.R.; Stoddart, J.F. Molecular-based electronically switchable tunnel junction

devices. J. Am. Chem. Soc. **2001**, *123* (50), 12632–12641.
12. Derycke, V.; Martel, R.; Appenzeller, J.; Avouris, Ph. Carbon nanotube inter- and intramolecular logic gates. Nano Lett. **2001**, *1* (9), 453–456.
13. Kwok, K.S.; Ellenbogen, J.C. Moletronics: Future electronics. Mater. Today **2002**, *5* (2), 28–37.
14. Cui, J.B.; Burghard, M.; Kern, K. Room temperature single electron transistor by local chemical modification of carbon nanotubes. Nano Lett. **2002**, *2* (2), 117–120.
15. Hong, B.H.; Bae, S.C.; Lee, C.-W.; Jeong, S.; Kim, K.S. Ultrathin single-crystalline silver nanowire arrays formed in an ambient solution phase. Science **2001**, *294*, 348–351.
16. Chen, Y.; Ohlberg, D.A.A.; Medeiros-Ribeiro, G.; Chang, Y.A.; Williams, R.S. Self-assembled growth of epitaxial erbium disilicide nanowires on silicon(001). Appl. Phys. Lett. **2000**, *76* (26), 4004–4006.
17. Iijima, S. Helical microtubules of graphitic carbon. Nature **1991**, *354*, 56–58.
18. Tans, S.J.; Devoret, M.H.; Dai, H.J.; Thess, A.; Smalley, R.E.; Geerligs, L.J.; Dekker, C. Individual single-wall carbon nanotubes as quantum wires. Nature **1997**, *386*, 474–477.
19. Wildöer, J.W.G.; Venema, L.C.; Rinzler, A.G.; Smalley, R.E.; Dekker, C. Electronic structure of atomically resolved carbon nanotubes. Nature **1998**, *391*, 59–62.
20. Braun, E.; Eichen, Y.; Sivan, U.; Ben-Yoseph, G. DNA-templated assembly and electrode attachment of a conducting silver wire. Nature **1998**, *391*, 775–778.
21. Fink, H.W.; Schönenberger, C. Electrical conduction through DNA molecules. Nature **1999**, *398*, 407–410.
22. Porath, D.; Bezryadin, A.; de Vries, S.; Dekker, C. Direct measurement of electrical transport through DNA molecules. Nature **2000**, *403*, 635–638.
23. Kasujov, A.Y.; Kociak, M.; Gueron, S.; Reulet, B.; Volkov, V.T.; Klinov, D.V.; Bouchiat, H. Proximity-induced superconductivity in DNA. Science **2001**, *291*, 280–282.
24. Yoo, K.-H.; Ha, D.H.; Lee, J.-O.; Park, J.W.; Kim, J.; Kim, J.J.; Lee, H.-Y.; Kawai, T.; Choi, H.Y. Electrical conduction through poly(dA)–poly(dT) and poly(dG)–poly(dC) DNA molecules. Phys. Rev. Lett. **2001**, *87* (19), 198102.
25. Bumm, L.A.; Arnold, J.J.; Cygan, M.T.; Dunbar, T.D.; Burgin, T.P.; Jones, L.; Allara, D.L.; Tour, J.M.; Weiss, P.S. Are single molecular wires conducting? Science **1996**, *271*, 1705–1707.
26. Reed, M.A.; Zhou, C.; Muller, C.J.; Burgin, T.P.; Tour, J.M. Conductance of a molecular junction. Science **1997**, *278*, 252–254.
27. Langlais, V.J.; Schlittler, R.R.; Tang, H.; Gourdon, A.; Joachim, C.; Gimzewski, J.K. Spatially resolved tunneling along a molecular wire. Phys. Rev. Lett. **1999**, *83* (14), 2809–2812.
28. Shimoura, T.; Akai, T.; Abe, T.; Ito, K. Atomic force microscopy observation of insulated molecular wire formed by conducting polymer and molecular nanotube. J. Chem. Phys. **2002**, *116* (5), 1753–1756.
29. Okawa, Y.; Aono, M. Nanoscale control of chain polymerization. Nature **2001**, *409*, 683–684.
30. Okawa, Y.; Aono, M. Linear chain polymerization initiated by a scanning tunneling microscope tip at designated positions. J. Chem. Phys. **2001**, *115* (5), 2317–2322.
31. Wegner, G. Topochemical polymerization of monomers with conjugated triple bonds. Makromol. Chem. **1972**, *154*, 35–48.
32. Tieke, B.; Lieser, G.; Wegner, G. Polymerization of diacetylenes in multilayers. J. Polym. Sci., Polym. Chem. Ed. **1979**, *17*, 1631–1644.
33. Rabe, J.P.; Buchholz, S.; Askadskaya, L. Scanning tunnelling microscopy of several alkylated molecular moieties in monolayers on graphite. Synth. Met. **1993**, *54*, 339–349.
34. Grim, P.C.M.; Feyter, S.D.; Gesquière, A.; Vanoppen, P.; Rücker, M.; Valiyaveettil, S.; Moessner, G.; Müllen, K.; De Schryver, F.C. Submolecularly resolved polymerization of diacetylene molecules on the graphite surface observed with scanning tunneling microscopy. Angew. Chem., Int. Ed. Engl. **1997**, *36* (23), 2601–2603.
35. Takami, T.; Ozaki, H.; Kasuga, M.; Tsuchiya, T.; Mazaki, Y.; Fukushi, D.; Ogawa, A.; Uda, M.; Aono, M. Periodic structure of a single sheet of a clothlike macromolecule (atomic cloth) studied by scanning tunneling microscopy. Angew. Chem., Int. Ed. Engl. **1997**, *36* (24), 2755–2757.
36. Hibino, M.; Sumi, A.; Hatta, I. Atomic images of saturated and unsaturated fatty acids at liquid / graphite interface and difference of tunneling currents between them observed by scanning tunneling microscopy. Jpn. J. Appl. Phys. **1995**, *34* (2A), 610–614. Part 1.
37. Groszek, A.J. Selective adsorption at graphite / hydrocarbon interfaces. Proc. R. Soc. Lond., A **1970**, *314*, 473–498.
38. McGonigal, G.C.; Bernhardt, R.H.; Thomson, D.J. Imaging alkane layers at the liquid / graphite interface with the scanning tunneling microscope. Appl. Phys. Lett. **1990**, *57* (1), 28–30.
39. Rabe, J.P.; Buchholz, S. Commensurability and

mobility in two-dimensional molecular patterns on graphite. Science **1991**, *253*, 424–427.
40. Venkataraman, B.; Breen, J.J.; Flynn, G.W. Scanning tunneling microscopy studies of solvent effects on the adsorption and mobility of triacontane/triacontanol molecules adsorbed on graphite. J. Phys. Chem. **1995**, *99* (17), 6608–6619.
41. Neumann, W.; Sixl, H. The mechanism of the low temperature polymerization reaction in diacetylene crystals. Chem. Phys. **1981**, *58*, 303–312.
42. Takabe, T.; Tanaka, M.; Tanaka, J. The electronic spectra of diphenylacetylene and 1,3-butadiyne. Bull. Chem. Soc. Jpn. **1974**, *47* (8), 1912–1916.
43. Bertault, M.; Fave, J.L.; Schott, M. The lowest triplet state of a diacetylene. Chem. Phys. Lett. **1979**, *62* (1), 161–165.
44. Albrecht, T.R.; Dovek, M.M.; Kirk, M.D.; Lang, C.A.; Quate, C.F.; Smith, D.P.E. Nanometer-scale hole formation on graphite using a scanning tunneling microscope. Appl. Phys. Lett. **1989**, *55* (17), 1727–1729.
45. Eigler, D.M.; Schweizer, E.K. Positioning single atoms with a scanning tunneling microscope. Nature **1990**, *344*, 524–526.
46. Kobayashi, A.; Grey, F.; Williams, R.S.; Aono, M. Formation of nanometer-scale grooves in silicon with a scanning tunneling microscope. Science **1993**, *259*, 1724–1726.
47. Avouris, Ph.; Walkup, R.E.; Rossi, A.R.; Akpati, H.C.; Nordlander, P.; Shen, T.-C.; Lyding, J.W.; Abeln, G.C. Breaking individual chemical bonds via STM-induced excitations. Surf. Sci. **1996**, *363* (1–3), 368–377.
48. Shen, T.C.; Wang, C.; Tucker, J.R. Al nucleation on monohydride and bare Si(001) surfaces: Atomic scale patterning. Phys. Rev. Lett. **1997**, *78* (7), 1271–1274.
49. Dujardin, G.; Walkup, R.E.; Avouris, Ph. Dissociation of individual molecules with electrons from the tip of a scanning tunneling microscope. Science **1992**, *255*, 1232–1235.
50. Stipe, B.C.; Rezaei, M.A.; Ho, W.; Gao, S.; Persson, M.; Lundqvist, B.I. Single-molecule dissociation by tunneling electrons. Phys. Rev. Lett. **1997**, *78* (23), 4410–4413.
51. Ma, L.P.; Yang, W.J.; Xie, S.S.; Pang, S.J. Ultrahigh density data storage from local polymerization by a scanning tunneling microscope. Appl. Phys. Lett. **1998**, *73* (22), 3303–3305.
52. Patitsas, S.N.; Lopinski, G.P.; Hul'ko, O.; Moffatt, D.J.; Wolkow, R.A. Current-induced organic molecule–silicon bond breaking: Consequences for molecular devices. Surf. Sci. Lett. **2000**, *457*, L425–L431.
53. Gao, H.J.; Sohlberg, K.; Xue, Z.Q.; Chen, H.Y.; Hou, S.M.; Ma, L.P.; Fang, X.W.; Pang, S.J.; Pennycook, S.J. Reversible, nanometer-scale conductance transitions in an organic complex. Phys. Rev. Lett. **2000**, *84* (8), 1780–1783.
54. Lopinski, G.P.; Wayner, D.D.M.; Wolkow, R.A. Self-directed growth of molecular nanostructures on silicon. Nature **2000**, *406*, 48–51.
55. Karpfen, A. Ab initio studies on polymers: IV. Polydiacetylenes. J. Phys., C. Solid State Phys. **1980**, *13*, 5673–5689.
56. Brédas, J.L.; Chance, R.R.; Silbey, R.; Nicolas, G.; Durand, Ph. A nonempirical effective Hamiltonian technique for polymers: Application to polyacetylene and polydiacetylene. J. Chem. Phys. **1981**, *75* (1), 255–267.
57. Salaneck, W.R.; Fahlman, M.; Lapersonne-Meyer, C.; Fave, J.-L.; Schott, M.; Lögdlund, M.; Brédas, J.L. Electronic structure of 4-BCMU polydiacetylene studied by angle-dependent photoelectron spectroscopy. Synth. Met. **1994**, *67*, 309–314.
58. Donovan, K.J.; Sudiwala, R.V.; Wilson, E.G. Fast photoconduction in Langmuir–Blodgett multilayers of polydiacetylenes. Thin Solid Films **1992**, *210/211*, 271–273.
59. Day, D.R.; Lando, J.B. Conduction in polydiacetylene bilayers. J. Appl. Polym. Sci. **1981**, *26*, 1605–1612.

Polymer–Clay Nanocomposites and Polymer Brushes from Clay Surfaces

P

Xiaowu Fan
Chuanjun Xia
University of Alabama at Birmingham, Birmingham, Alabama, U.S.A.

Rigoberto C. Advincula
University of Alabama at Birmingham, Birmingham, Alabama, U.S.A.
University of Houston, Houston, Texas, U.S.A.

INTRODUCTION

This paper summarizes the recent advances in the synthesis of polymer–clay nanocomposite materials and polymer brushes on clay surfaces with the focus on the in situ polymerization method. This new type of composite materials is based on a wide range of polymers as the matrix and layered silicate clays as the reinforcing filler. Its advantages, structures, and synthesis routes as well as the special properties of layered silicate are briefly introduced. Among various synthesis methods including exfoliation–adsorption, melt intercalation, and in situ polymerization, detailed discussions are focused on the in situ polymerization approach, which is then categorized by the polymerization mechanism employed. A comprehensive coverage is given to the in situ free radical polymerization method, including free radical polymerization with original clay, organoclay modified with nonreactive ammonium cations or polymerizable surfactants, and clay intercalated with cationic initiator derivatives, i.e., the surface-initiated polymerization strategy. As different strategies applied by various studies are compared, the emphasis of the discussion is placed on the synthesis scheme and the resulting structure of the nanocomposites. Short discussions on polymer brushes and in situ living polymerization methods on clay particle surfaces are offered because of the limited number of related publications. From the standpoint of the category of polymeric materials, this review is primarily based on thermoplastic polymers that can be produced by polymerizing their vinyl monomers. Other materials from other in situ polymerization mechanisms, such as ring opening polymerization, are briefly introduced at the end of this paper.

BACKGROUND

Polymer–clay nanocomposite (PCN) materials fall into the category of organic–inorganic hybrids consisting of organic polymers as the matrix and inorganic clay minerals as the reinforcing filler. The unique characteristics of PCN materials lie in the nanometer-scale mixing of the two components and the molecular-level interactions between them. They have a wide variety of highly improved mechanical, thermal, and barrier properties as compared with those of pure polymers and their macroscale counterparts, which was demonstrated in the seminal work on nylon 6-clay hybrids (NCHs) by Toyota researchers in the early 1990s (Table 1).[1–4]

Properties of Layered Silicates

Polymer–clay nanocomposite materials are also usually referred to as polymer-layered silicate (PLS) nanocomposites because the inorganic fillers used are mostly natural or synthetic silicate materials such as montmorillonite, hectorite, laponite, fluorohectorite, etc. They possess a unique layered morphology of ordered stacks of quasi two-dimensional silicate sheets whose dimensions are about 1 nm in thickness and a few hundred nanometers in width and length. As illustrated in Fig. 1, the crystallographic structure in one single layer consists of an edge-shared octahedral sheet of alumina or magnesia sandwiched between two tetrahedral silica sheets. Metal ions such as Na^+, Li^+, and Ca^{2+} usually occupy the interlayer spacings/galleries to counterbalance the negative charges in the anionic lattice, which are generated by the isomorphous cation substitution within the layers.[5] The

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013459

Table 1 Property comparison of NCH, NCC, and pure nylon 6 as presented in Ref. [1]

Specimen (montmorillonite wt.%)	Tensile strength (MPa)	Charpy impact strength (MPa)	HDT[a] at 18.5 kg/cm^2 (C)	Rate of water adsorption at 23°C, 1 day (%)
NCH[b] (4.2)	81	2.8	145	0.51
NCC[c] (5.0)	64	2.2	89	0.90
Nylon 6 (0)	72	2.3	65	0.87

[a]Heat distortion temperature (HDT).
[b]Nanoscale nylon 6-clay hybrid (NCH).
[c]Macroscale nylon 6-clay composite (NCC).

metal ions inside the silicate layers can be replaced by various organic cations such as alkylammonium by the cation exchange process, forming the "organically modified" layered silicate (OLS, also referred to as organoclay or organophilic clay). The exchangeable cations and layered morphology of these silicate clays lead to their two important physical properties: cation exchange capacity (CEC, often measured in the unit of meq/100 g, i.e., milliequivalents per 100 g) and specific surface area. For example, the specific area of montmorillonite clay has a range of 700–800 m^2/g and its CEC varies from 80 to 150 meq/100 g.

Structures of Polymer–Clay Nanocomposite Materials

Beyond traditional phase-separated macroscale composites, PCN materials can be structurally characterized into two categories in terms of their different microstructures resulting from different degrees of homogeneity between the organic and inorganic components:[6,7] 1) *Intercalated*, in which extended polymer chains are inserted between the clay layers, leading to an ordered multilayer morphology of alternating polymer/clay layers. As a result, the *d* spacing of the clay particle stacks is usually increased; 2) *exfoliated/delaminated*, in which the silicate layers are exfoliated into discrete sheets and uniformly dispersed in a continuous polymer matrix. This structure gives rise to a complete loss of the structural registry of the clay layers. Because of greater phase homogeneity with the latter, the exfoliated structure is more effective in improving the properties of the polymer, especially at very low clay loading. X-ray diffraction (XRD) *d*-spacing measurement and transmission electron microscopy (TEM) observation are two commonly used and complementary spectroscopic and microscopic tools to characterize and differentiate these microstructures. In addition to the two well-defined structures, intermediate structures comprising both intercalated and exfoliated domains are also observed. For the PCN samples of this mixed structure, the XRD peaks are often broadened and TEM analysis is imperative to characterize the overall structure.[8]

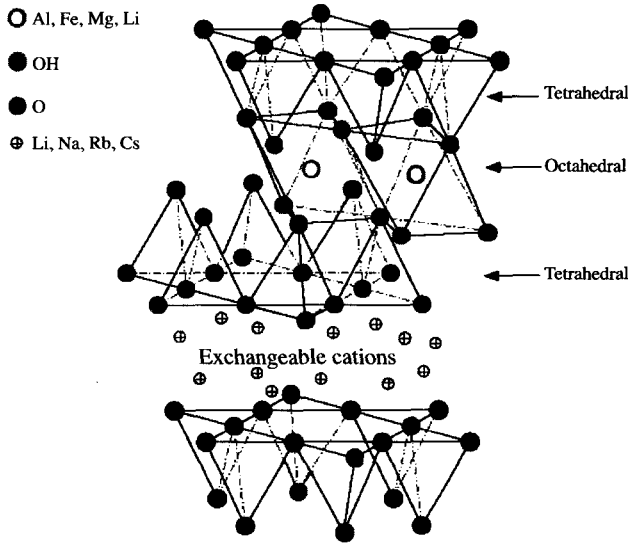


Fig. 1 Illustration of the structure of layered silicates. (From Ref. [5]. Copyright 2003 Springer-Verlag GmbH & Co.KG.) *(View this art in color at www.dekker.com.)*

Preparation Methods of Polymer–Clay Nanocomposite Materials

There are different numbers of synthetic routes toward PCN materials as classified by previous review articles.[9,10] The three primary methods are as follows:

Exfoliation–adsorption: This generally involves the exfoliation of clay in a proper solvent and the adsorption of a soluble polymer to the clay surfaces. Polymer–clay nanocomposite materials, especially those based on polymers with high hydrophilicity, such as poly(vinyl alcohol), poly(ethylene oxide), and poly(acrylic acid), can be produced.[11,12] However, the main drawback is that a suitable solvent to promote compatibility of

the inorganic clay and organic polymers is not always available.

Melt intercalation: This is a simple mix-and-heat process, in which the layered silicate, usually organically modified, is mixed with the polymer matrix in the molten state. As compared with exfoliation–adsorption method, it is a versatile and green approach as a wide variety of polymers can be applied and no solvents are needed.[13–16] It also provides a model system in which the behavior of confined polymers can be studied.[17–22] Nonetheless, the ideal exfoliated structure is not always achieved.

In situ polymerization: Monomers/comonomers are polymerized with the presence of original/modified clay via different polymerization mechanisms. This is also often referred to as intercalative or interlamellar polymerization, in which the chain growth is expected to be inside the galleries of the layered silicate. The advantage of this method is that as the polymerization inside the galleries progresses, the clay particles will be gradually swollen and ultimately exfoliated into discrete layers. As a result, this approach can more likely lead to PCN materials with a delaminated structure.

In this article, we would like to focus on the strategy of in situ polymerization toward polymer brushes on clay particle surfaces and PCN materials. Details and characteristics of other methods such as template synthesis[23] and even a "general approach"[24] to PCN materials can be found in their respective reviews/reports and will not be discussed in this article.

Polymer Brushes on Clay Surfaces

Polymer brushes are defined as an assembly of polymer chains that are immobilized at one end to a surface by tethering, which is sufficiently dense so that the chains are forced to stretch away from the surface as a result of interchain repulsion.[25] Because of the layered morphology and negative surface charges, polymer chains that are ionically bound to clay surfaces can also be considered as polymer brushes if the surface charge density of the clay substrate is high enough so that the average distance between anionic sites is smaller than the radius of gyration of a polymer. They can exist in two different geometries. One is that the polymer brushes can be formed in bulk either by physisorption of one-end-charged polymers as described in the *exfoliation–adsorption* method (grafting-to) or by in situ polymerization using cationic monomer/initiator derivatives (grafting-from). Another is that the brushes can be grafted to/from a flat layer of clay particles on planar substrates prepared by the self-assembled monolayer (SAM) technique or the alternate polyelectrolyte deposition (APD) process.[26–28] The former case has been extensively investigated in the form of PCN materials while there are few studies in the literature about the latter.

P

METHODS OF IN SITU POLYMERIZATION

The history of in situ polymerizations with clay can be traced back to the 1960s. The type of clay involved was original clay or organoclay modified with organic cations with various functionalities. The form of polymerization ranged from bulk, solution, suspension, and emulsion. The PCN products could be thermoplastics or thermosets. Here we categorize the following in situ methods based on their individual polymerization mechanism. The emphasis of the discussion is focused on the synthetic routes and structures of the PCN materials rather than to their practical properties. It should also be noted that there is another case of in situ polymerization, in which clay minerals, especially when exchanged with transition metal cations, can also initiate/catalyze the polymerization of different monomers;[29,30] but this will not be discussed here.

In Situ Free Radical Polymerization

Many important thermoplastic polymers, such as polystyrene (PS) and poly(methyl methacrylate) (PMMA), can be produced by free radical polymerization mechanism using their vinyl monomers. The following discussion is focused on the PCN materials based on these two typical polymers, and in situ free radical methods are subcategorized by the initiators used.

Intercalative free radical polymerization using traditional initiators

Early studies on in situ free radical polymerization were focused on mixing vinyl monomers with unmodified/modified clay followed by polymerization with traditional azo, peroxide, or persulfate initiators. In the mid-1960s, in situ free radical polymerization with pristine clays was demonstrated by Blumstein[31] and Solomon and Rosser.[32] The effects of different types of clay on polymerization mechanisms and the interactions between different monomers and clay were studied. Lee and Jang[33] obtained PMMA–clay materials of intercalated structure through emulsion polymerization with original clay using potassium persulfate as the initiator. Later, they used the same method to prepare PS–clay nanocomposites. As compared with their earlier PMMA materials, the

deceased d spacings of the intercalated PS composites were attributed to stronger hydrophobicity of the styrene monomer, which had less compatibility to penetrate into the galleries of the hydrophilic clay.[34] Exfoliated PMMA nanocomposites was accomplished by Bandyopadhyay et al.[35] using different clays (montmorillonite and fluorohectorite) but similar initiator and emulsion methods. Using the same initiator, suspension polymerizations of MMA initiated with the presence of montmorillonite clay were investigated by Al-Esaimi,[36] but structural information of the composites was not reported. Structures from direct PMMA melt intercalation and in situ polymerization with alkylammonium-modified clay and 2,2′-azobisisobutyronitrile (AIBN) initiator were compared by Tabtiang et al.[37] and a finer degree of clay dispersion was observed for the latter. Exfoliated PS–clay nanocomposites were achieved by Chen et al.[38] using emulsion polymerization with ammonium persulfate as the initiator and cetyltrimethylammonium as the organic modifier of the clay. Similarly, Okamoto et al.[39] synthesized partially exfoliated PMMA–clay materials and intercalated PS–clay composites by bulk polymerization with lipophilized smectic clays. Later, they studied the effects of the addition of a small amount of different polar comonomers (acrylate and acrylamide derivatives) on the structure and properties of the PMMA–clay composites.[40] Likewise, Dietsche et al.[41] found that the addition of a small amount of dodecylmethacrylate (LMA) comonomer as a compatibilizer with organoclay [modified by dioctadecyldimethylammonium (DOD) cation] could promote clay exfoliation in the poly(MMA-co-LMA) matrix. By bulk polymerizations with AIBN as the initiator, Doh and Cho[42] systematically investigated the effect of different organic modifiers on the structure and properties of PS–clay nanocomposites. Although exfoliated structure was not observed, one of the three PS–clay products exhibited better clay dispersion and thermal property. The authors ascribed it to the structural affinity between the styrene monomer and the benzyl unit in one of the three alkylammonium cations employed. From the discussions above, it can be concluded that in the in situ intercalative polymerization, the compatibility between the monomer/comonomer and organic groups in the organoclay plays an important role in the structure and properties of PCN materials. Fig. 2 illustrates the preparation method used by Chen et al.,[38] which can also be used to schematically generalize the concept of in situ intercalative polymerization with organoclay modified by ammonium cations. Other than homopolymers, PCNs based on copolymers such as poly(styrene-co-acrylonitrile) (SAN) were also prepared by in situ polymerization with pristine or modified clay, indicating the versatility of this strategy.[43,44]

In the studies cited above, the polymer chains are freely formed and are not bound to the clay surfaces, as can be seen in Fig. 2. The following discussion will introduce recent studies on the in situ polymerization with clay modified by polymerizable cations by which a portion of the polymer chains is end-tethered to the clay surfaces. With two kinds of modified clays functionalized by charged acrylate derivatives, Biasci et al.[45] achieved intercalated PMMA–clay nanocomposites by solution polymerization initiated by AIBN. Similarly and using the same initiator, intercalated PS–clay materials were synthesized by Akelah and Moet[46,47] through ion exchanging with a vinylbenzylammonium cation and subsequent in situ free radical polymerization of styrene. As a continuation of the previous work, Dietsche et al.[41] synthesized the poly (MMA-co-LMA) materials from organoclay modified by a reactive cation containing two MMA end units [bis(2-methacryloyloxyethyl-)methyl-n-dodecyl-ammonium (DMMA)]. As compared with the old PCNs from nonreactive cation DOD, unusual combination of high stiffness and high toughness was observed in the new nanocomposites and attributed to the surface attachment of the polymers through covalent bonding with the polymerizable DMMA.[48] Exfoliated PS–clay nanocomposites were

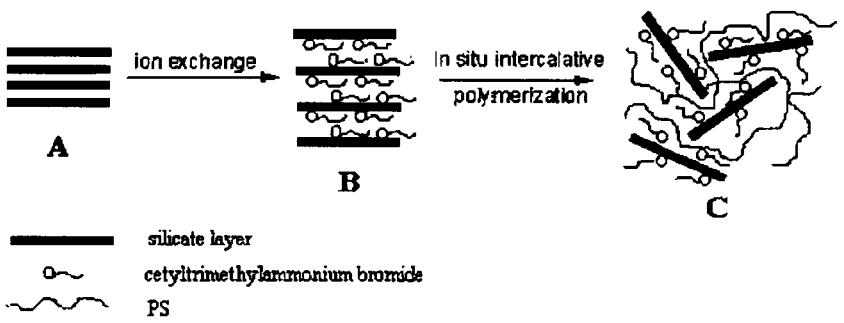


Fig. 2 Schematic illustration of the preparation of PS–montmorillonite nanocomposite via in situ intercalative polymerization. (A) Original sodium montmorillonite. (B) Cetyltrimethylammonium bromide (CTAB) exchanged montmorillonite. (C) Exfoliated PS–montmorillonite nanocomposite. (From Ref. [38]. Copyright 2003 Acta Materialia and Elsevier Science Inc. Ltd.)

achieved from in situ bulk polymerization by Fu and Qutubuddin.[49] In their studies, organophilic clay was prepared by the insertion of a polymerizable cationic surfactant vinylbenzyldimethyldodecylammonium chloride (VDAC), which was then directly dispersed in monomer solution followed by initiation by AIBN. They believed that exfoliation resulted from the special properties of the VDAC modifier: the reactivity of the vinylbenzyl group, the solubility in styrene monomer, and the better swelling ability for clay.[50] Their results were confirmed by Zhu et al.,[51] as essentially the same reactive modifier also resulted in exfoliated PS–clay nanocomposites. In their comparative studies using bulk polymerization and AIBN initiator but different cationic surfactants, better thermal stability was found in the organoclay modified by a phosphonium salt than in the other two modified by ammonium salts, which might be useful if high processing temperatures were required for the PCN materials. Similar results were recently observed by Yeh et al.[52] as they found the PMMA–clay nanocomposite from alkylphosphonium as the intercalating agent displayed higher glass transition temperature (T_g) than that from conventional alkylammonium salt. Depending on the clay content, exfoliated, intercalated, and mixed intercalated–exfoliated structures were found in their PMMA materials via in situ solution polymerization with benzoyl peroxide (BPO) as the initiator. However, it should be noted that neither of the "onium" molecules contained polymerizable groups. Clay exfoliation in both PS–[53] and PMMA–clay[54] nanocomposites was accomplished by Choi, Kim, and their coworkers through emulsion polymerizations with 2-acrylamido-2-methyl-1-propane-sulfonic acid (AMPS) as the reactive surfactant. They were able to obtain the end-tethered polymers by a reverse ion-exchange process. The copolymerization between AMPS and styrene was confirmed by NMR measurements. Likewise, Zeng and Lee[55] prepared exfoliated PS– and PMMA–clay materials by bulk polymerization with clay modified by an ammonium salt containing an MMA end unit. Furthermore, by investigating the effects of different initiators including BPO and AIBN, they also stressed that compatibility of the monomer, initiator, and organoclay could strongly influence the clay dispersion in the polymer matrix. Based on PS and PMMA, Wang et al.[56] comprehensively studied the methods for PCN materials including in situ bulk, solution, suspension, and emulsion polymerization as well as melt intercalation. The effects of different initiators and ammonium ions on the in situ polymerizations were also compared. They also found that clay exfoliation were improved by the polymerizable group in the organic modifier that could participate in the polymerization. Fig. 3 shows the synthetic route utilized by Akelah and Moet,[46] representing the scheme of tethering polymer chain to clay surfaces with cationic comonomer derivatives. Here the end-tethered chains can also be considered as polymer brushes formed in-bulk by the grafting-from approach.

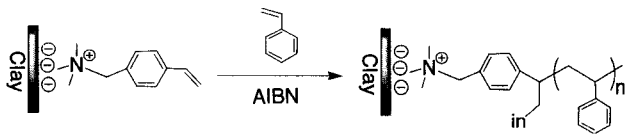


Fig. 3 In situ polymerization scheme as employed in Ref. [46]. The free radical polymerization of styrene with clay modified by a cationic comonomer resulted in surface-bound polymer chains.

From the previous discussions, it is obvious that surface-attached polymer chains through polymerizable modifiers can indeed improve clay particle exfoliation as compared with nonreactive alkylammoniums. Besides PS and PMMA, this concept has also been demonstrated by other PCNs such as poly(acrylamide)[57] and poly(acrylonitrile)[58] via in situ polymerization using traditional free radical initiators.

It is worth mentioning that the coworkers of Dietsche further continued their previous studies[41,48] by utilizing protonated amine-terminated oligostyrene[59] and cyclic amidinium cation-terminated oligoMMA and oligostyrene[60] as organic modifiers of the clay, which was applied in the subsequent melt compounding/extrusion processes with different polymers including polystyrene, PMMA, and styrene–acrylonitrile copolymer. The cationic oligomers (number average molecular weight 3000 g/mol for amidine-terminated species and 5800 g/mol for amine-terminated species) were first synthesized by controlled free radical and anionic polymerization methods and then attached to clay surfaces by ion exchange processes. Unlike clays intercalated with various alkylammonium cations, XRD measurements were even unable to detect the significantly increased d spacings of the well-swollen organoclays inserted with the oligomeric cations. It has to be pointed out that the OLSs here are analogs to the PCNs prepared by the exfoliation–adsorption method, except that the organic components are oligomeric cations rather than water-soluble polymers. Furthermore, these end-tethered oligomers are similar to the polymer brushes on clay surfaces by the grafting-to strategy described before, except for their low molecular weight nature. Effective exfoliation was observed by scanning electron microscopy (SEM) of the PCN materials after melt intercalation. Interesting rheological and thermal properties were found to be related with the interactions among polymer matrix, oligomer brushes tethered on clay surfaces, and the clay particles themselves.[61] Citing amine-terminated oligostyrene by anionic polymerization as an example, their synthetic scheme is illustrated in Fig. 4.

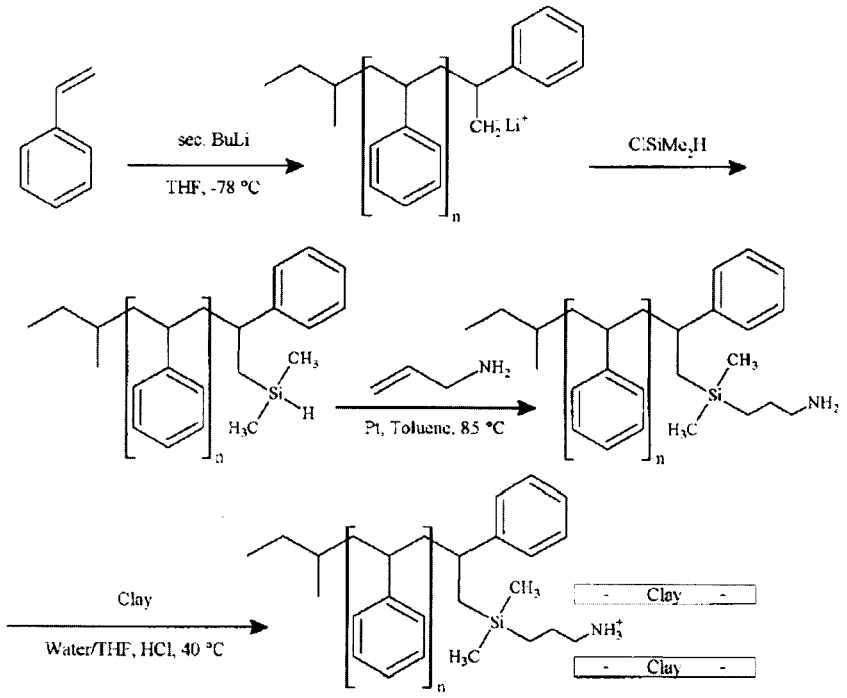


Fig. 4 Synthesis of the oligostyrene and preparation of the organoclay. (From Ref. [59]. Copyright 2003 Wiley-VCH Verlag GmbH.)

Surface-initiated free radical polymerization using cationic initiator derivatives

As for in situ free radical polymerization with clay, polymerization initiator derivatives can also be inserted into clay particles' galleries instead of reactive monomer cations or nonreactive surfactants. In this case, the surface-attached initiator not only renders the clay organophilic but also provides the initiation functionality for the subsequent in situ polymerization. Because it is the surface-attached initiator that activates the polymerization process and the polymer chain actually propagates from the surface, this type of in situ polymerization is also called the surface-initiated polymerization (SIP). The advantage of in situ polymerization with the initiator-intercalated clay lies in a sound assumption that, as the bound polymer chains propagate, the ordered silicate layers can be gradually pushed apart and finally exfoliated in the polymer matrix. Furthermore, as compared with the polymerizable surfactant route discussed above, the SIP strategy can presumably generate more end-tethered polymer chains with greater length, which provide more efficiency in exfoliation. Free radical SIPs using different AIBN initiator variants were demonstrated on various surfaces including spherical silica gel particles,[62,63] planar SiO_x[64] and gold substrates,[65,66] etc., on which the initiators were immobilized using the SAM method through covalent bonding. As for SIPs from clay surfaces, the initiator must be rendered cationic for ionic bonding with clay surfaces. The bound polymer chains after SIP can be cleaved for analysis by a counter-ion-exchange process using an excessive amount of alkali compound with greater cationic strength such as lithium bromide.

2,2′-Azobis(isobutyramidine hydrochloride) (AIBA) is a water-soluble, commercially available, and bicationic azo compound that can be used as a surface-attached free radical initiator. In the mid-1960s, its feasibility was demonstrated by Dekking[67,68] through the SIPs of several vinyl monomers. By bulk and emulsion polymerization, the effects of different clays (kaolin, bentonite, and hectorite) on the kinetics of the attached initiator decomposition and the subsequent chain propagation were studied. Similarly, Meier et al.[69] also employed AIBA to polymerize styrene from high-surface-area mica powder and the SIP mechanism was investigated. Although high molecular weight grafted polymer was obtained, they argued that the surface initiation and chain propagation played a minor role in the formation of surface-attached polymers. Instead, bound chains were created by the reaction between free polymers formed in solution and the surface-bound products of AIBA decomposition. However, no structural information of the composite products was reported. Huang and Brittain,[70] employing the same AIBA initiator and suspension SIP, achieved PMMA–clay nanocomposites with exfoliated structure. For comparison, they also prepared PMMA materials

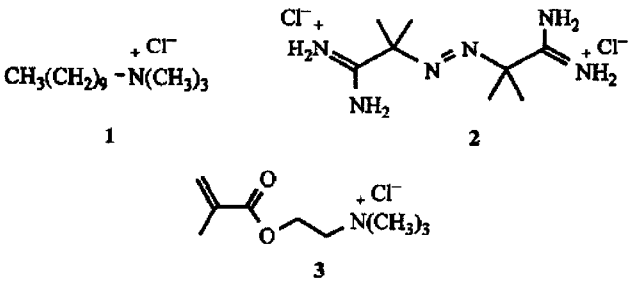


Fig. 5 Structures of the three cations used to modify layered silicates. 1) An alkyl ammonium chloride salt; 2) the AIBA initiator; 3) a reactive surfactant containing a polymerizable MMA unit. (From Ref. [70]. Copyright 2003 American Chemical Society.)

based on nonreactive and polymerizable methacrylammonium-modified clays via in situ polymerization using traditional AIBN initiator. It was observed that exfoliated structure was achieved and preserved after melt processing for the nanocomposites from AIBA and methacrylammonium-modified clays. In contrast, alkylammonium organoclay yielded only the intercalated structure both before and after melt processing and an even greater degree of order after melt processing. Thus, the authors believed that the end-tethered polymer chains not only promoted clay particle exfoliation but also helped preserve the exfoliated structure at elevated temperature by hindering the aggregation of the silicate layers. Between the nanocomposites from SIP with AIBA and polymerizable cation with AIBN, better thermal stability was observed in the former. Fig. 5 shows the structures of their three cationic modifiers, which also clearly present the three foregoing concepts in OLS modifications.

We synthesized another water-soluble monocationic azo compound.[71] The advantage of this AIBN derivative is that, unlike AIBA, the positive charge at one end is pH independent, which can tolerate different polymerization conditions. In addition, the other alkyl end of the molecule can greatly enhance the homogeneity of clay–monomer/solvent mixture of the SIP system. It should also be pointed out that, other than bound polymers, free ones also form in bulk/solution resulting from the other unbound free radical after the decomposition of the azo group. Through solution SIP, this initiator was first used to graft polystyrene brushes from clay surfaces adsorbed by the SAM technique on planar substrates. Fig. 6 shows the structure of the initiator as well as the synthesis scheme of the polymer brush and the free polymer formation. By surface-sensitive characterizations including ellipsometry and surface plasmon spectroscopy (SPS), it was found that the average initiator graft density (on the order of 0.1–1 molecules/nm^2) and polymer layer thickness on flat clay surfaces were much lower than those on flat SiO_x substrates from which the initiator/polymer were grafted by covalent bonding (on the order of 2–3 molecules/nm^2

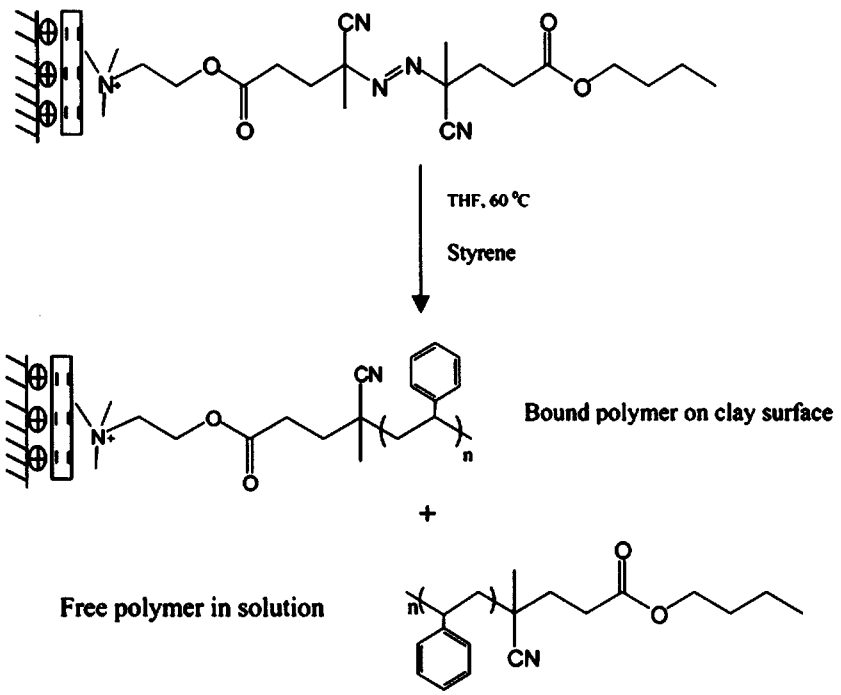


Fig. 6 Illustration of polymer brushes grafted from clay particles absorbed on flat substrates and free polymer formation. (From Ref. [71]. Copyright 2003 American Chemical Society.)

based on silane surface coupling).[72] The reason was believed to be that the charge density of clay surfaces was much lower than the density of the reactive sites of the SiO_x substrates. Therefore the bound polymers grafted from clay particles existed in the morphology of "mushrooms" rather than "brushes." It should be noted that the results depend on the physical properties of a specific type of clay, i.e., its CEC and specific surface area. The graft density can be varied if another type of clay is applied.

We further synthesized another similar but bicationic initiator, which was compared with the monocationic one in terms of the structures of their modified clays and the polystyrene–clay nanocomposites from solution SIPs.[73] The structure of intercalated clays after ion exchange was first studied by XRD measurements. Despite the similar structure and chain length of the two molecules, lower *d* spacing and less degree of order were observed in the organoclay modified by the bicationic initiator. This result was attributed to the combination of the two attachment possibilities of the bicationic initiator: 1) The two cationic ends were attached to the same side of one clay particle and; 2) they were bound to two sides of two adjacent particles. As for the SIP products, it was found that the monocationic initiator-modified clay was more likely to produce exfoliated structure. It was also found that the molecular weights of bound polymers were lower than that of the free ones. The differences caused by the two initiators, including organophilicity of the modified clays, clay particle dispersibility during SIP, diffusion and addition kinetics of monomer, and chain transfer and termination possibilities, were discussed and compared.

Using the monocationic initiator, we also prepared PMMA–clay nanocomposites via bulk, suspension, and solution SIPs.[74] Exfoliated structure was achieved by solution and bulk methods while suspension polymerization yielded a mixed exfoliation–intercalation morphology. It was ascribed to the amphiphilic characteristic of the initiator-modified clay, which probably formed aggregation at the monomer–water interface during suspension SIP. Intriguing results such as the effect of the affinity between immobilized initiator and monomer on the product structure, molecular weights, and molecular weight distributions of free and bound polymers as shown by gel permeation chromatography (GPC) analysis, were also observed and discussed.[75] It should be noted that although the molecular weight of the bound polymer was lower than that of the free one, it was still much higher as compared with that of the bound chains from the reactive surfactant approach, where only oligomers were obtained.[54] This proves the higher efficiency of producing long end-tethered chains in SIP methods than in the polymerizable modifier strategy. Fig. 6 can also be used to represent the synthesis scheme used here except that the monomer is MMA and the clay particle is not ionically adsorbed to a planar substrate.

In Situ Living Polymerization

It is well known that living polymerization methods are capable of synthesizing homopolymers with narrow polydispersity and controllable molecular weight and copolymers with desired architecture. On the other hand, they generally require more demanding polymerization conditions than the ordinary free radical approach. As compared with the intensive studies on living polymerization from other surfaces with various geometry (Ref. [76] and references therein), there were just a few attempts to prepare polystyrene–clay materials through in situ surface-initiated living polymerization from clay surfaces.

Living anionic surface initiated polymerization

We synthesized a cationic derivative of 1,1-diphenylethylene (DPE) to anionically polymerize styrene from clay surfaces.[77,78] 1,1-Diphenylethylene was first modified to be a water-soluble ammonium bromide salt and then intercalated into montmorillonite clay by ion exchange.

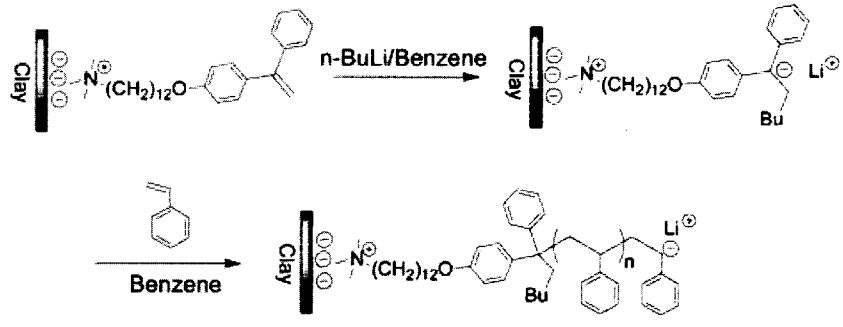


Fig. 7 Schematic illustration of living anionic surface-initiated polymerization (LASIP) from layered clay particles. (From Ref. [78]. Copyright 2003 American Chemical Society.)

The subsequent living anionic surface-initiated polymerization (LASIP) successfully yielded free and bound polystyrene with controllable molecular weight and narrow molecular weight distribution. However, the bound polymers had lower molecular weight than the free ones and only intercalated structure was obtained in the composite products. The authors believed that the dispersibility of intercalated clay and monomer diffusion kinetics in this heterogeneous polymerization system were critical factors of the LASIP process. Fig. 7 shows the synthetic route.

Living free radical polymerization

Exfoliated polystyrene–clay nanocomposites with controllable molecular weight were prepared by Weimer et al.[79] by intercalating an ammonium cation initiator containing a nitroxide moiety into the montmorillonite clay. The following living free radical polymerization (LFRP) protected by the nitroxyl group effectively produced polystyrene materials with narrow polydispersity and exfoliated structure. Moreover, while keeping the clay loading constant, they were able to control the molecular weight of the polymer matrix by varying the initiator concentration by cointercalating a noninitiating ammonium salt. The livingness of the reaction and block copolymerization viability was further confirmed by a successful effort of chain extension. It is worth mentioning that this is the first time that an exfoliated PCN material with controllable molecular weight has been achieved based on in situ polymerization with a vinyl monomer. Fig. 8 illustrates the synthesis scheme.

Other In Situ Polymerization Methods

Other than thermoplastic polymers prepared by in situ addition polymerization of their vinyl monomers, there has been tremendous work on other polymeric mateials by other polymerization mechanisms. In fact, early studies by Toyota scientists on nylon-based PCN materials demonstrated in situ ring opening polymerization of ε-caprolactam with organoclays modified by protonated amino acids.[80–82] By using the similar strategy, Messersmith and Giannelis[83,84] synthesized poly(ε-caprolactone)–clay nanocomposites. Aside from their highly improved mechanical, thermal, and barrier properties, unique viscoelastic behaviors of those end-tethered polymers were also discovered.[85,86] The latest progress on in situ ring opening polymerization can be found in the recent reports and related reviews by Dubois et al.[87–89] Indeed, the versatility of the in situ polymerization strategy has been demonstrated by a broad spectrum of other polymers via their individual polymerization mechanism including thermosets,[90–93] elastomers,[94,95] polyolefins,[96,97] and even conducting polymers.[98–100]

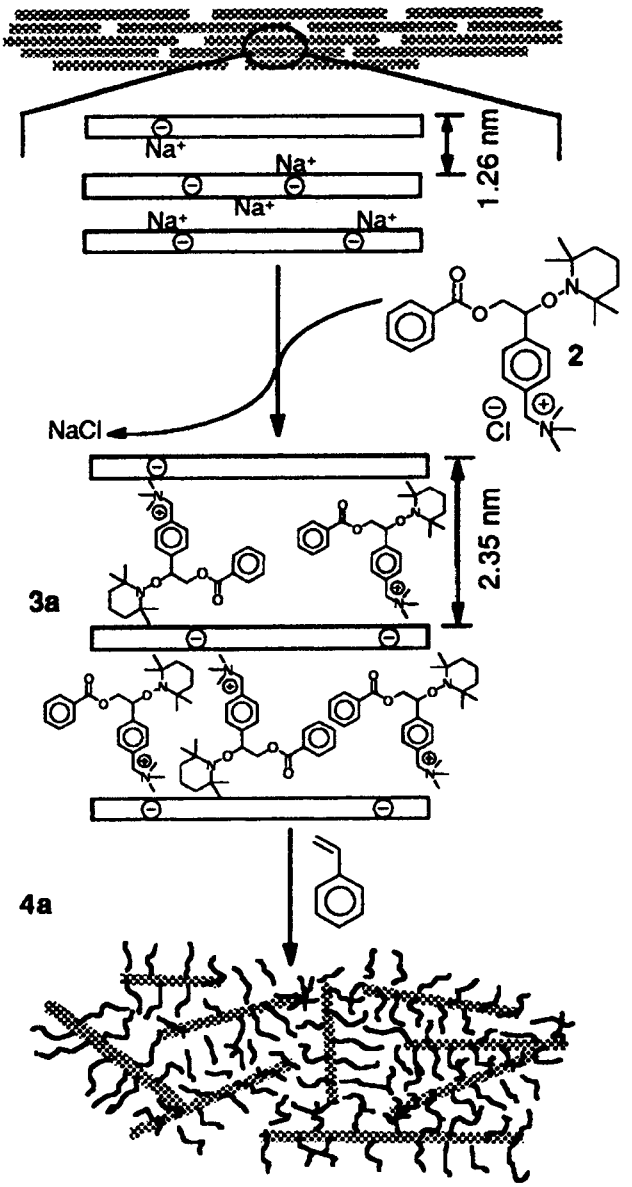


Fig. 8 Exfoliated polystyrene–clay nanocomposites by living free radical polymerization. (From Ref. [79]. Copyright 2003 American Chemical Society.)

CONCLUSION

In situ polymerization synthesis methods for PCN materials and polymer brushes on clay nanoparticles are summarized and categorized. The emphasis is placed on free radical and living polymerization approaches. Earlier investigations have shown that in situ polymerization with organoclays modified by reactive modifiers, in which a portion of the polymers is attached to clay surfaces by polymerizable cations, is more likely to give exfoliated

PCNs than with organoclays modified by nonreactive alkylammoniums or pristine clay. Previous studies have also demonstrated that in situ polymerization by SIP method, in which more and longer polymer chains can be end-tethered to clay surfaces by cationic initiators than by polymerizable modifiers, can also achieve effective exfoliation. The viability of achieving molecular weight-controllable PCNs has been demonstrated as well by several attempts on in situ living polymerization approaches from clay surfaces. Although the discussions in this paper are primarily based on thermoplastic polymers from addition polymerization of their vinyl monomers, the versatility and efficiency of in situ polymerization strategy have been verified by PCNs via other in situ polymerization mechanisms, which are briefly introduced. The related report and review references are also given.

ACKNOWLEDGMENT

The authors would like to thank the U.S. Army Research Office, Alabama Tri-Campus Materials Science Program, and the Chemistry Department of the University of Alabama at Birmingham for financial support.

REFERENCES

1. Okada, A.; Kawasumi, M.; Kurauchi, T.; Kamigaito, O. Synthesis and characterization of a nylon 6-clay hybrid. Polym. Prepr. **1987**, *28*, 447–448.
2. Usuki, A.; Kojima, Y.; Kawasumi, M.; Okada, A.; Fukushima, Y.; Kurauchi, T.; Kamigaito, O. Synthesis of nylon 6-clay hybrid. J. Mater. Res. **1993**, *8* (5), 1179–1184.
3. Kojima, Y.; Usuki, A.; Kawasumi, M.; Okada, A.; Fukushima, Y.; Kurauchi, T.; Kamigaito, O. Mechanical properties of nylon 6-clay hybrid. J. Mater. Res. **1993**, *8* (5), 1185–1189.
4. Kojima, Y.; Usuki, A.; Kawasumi, M.; Okada, A.; Kurauchi, T.; Kamigaito, O. Sorption of water in nylon-6 clay hybrid. J. Appl. Polym. Sci. **1993**, *49* (7), 1259–1264.
5. Giannelis, E.P.; Krishnamoorti, R.; Manias, E. Polymer-silicate nanocomposites: Model systems for confined polymer and polymer brushes. Adv. Polym. Sci. **1999**, *138*, 108–147.
6. Burnside, S.D.; Giannelis, E.P. Synthesis and properties of new poly(dimethylsiloxane) nanocomposites. Chem. Mater. **1995**, *7* (9), 1597–1600.
7. Lan, T.; Pinnavaia, T.J. Clay-reinforced epoxy nanocomposites. Chem. Mater. **1994**, *6* (12), 2216–2219.
8. Huang, X.; Lewis, S.; Brittain, W.J.; Vaia, R.A. Synthesis of polycarbonate-layered silicate nanocomposites via cyclic oligomers. Macromolecules **2000**, *33* (6), 2000–2004.
9. Alexandre, M.; Dubois, P. Polymer-layered silicate nanocomposites: Preparation, properties and uses of a new class of materials. Mater. Sci. Eng., R **2000**, *28* (1–2), 1–63.
10. Lagaly, G. Introduction: From clay mineral–polymer interactions to clay mineral–polymer nanocomposites. Appl. Clay Sci. **1999**, *15* (1–2), 1–9.
11. Ogata, N.; Kawakage, S.; Ogihara, T. Poly(vinyl alcohol)–clay and poly(ethylene oxide)–clay blend prepared using water as solvent. J. Appl. Polym. Sci. **1997**, *66* (3), 573–581.
12. Billingham, J.; Breen, C.; Yarwood, J. Adsorption of polyamine, polyacrylic acid, and polyethylene glycol on montmorillonite: An in situ study using ATR–FTIR. Vibr. Spectrosc. **1997**, *14* (1), 19–34.
13. Vaia, R.A.; Ishii, H.; Giannelis, E.P. Synthesis and properties of two-dimensional nanostructures by direct intercalation of polymer melts in layered silicates. Chem. Mater. **1993**, *5* (12), 1694–1696.
14. Giannelis, E.P. Polymer layered silicate nanocomposite. Adv. Mater. **1996**, *8* (1), 29–35.
15. Kawasumi, M.; Hasegawa, N.; Kato, M.; Usuki, A.; Okada, A. Preparation and mechanical properties of polypropylene–clay hybrids. Macromolecules **1997**, *30* (20), 6333–6338.
16. Liu, L.; Qi, Z.; Zhu, X. Studies on nylon-6 clay nanocomposites by melt-intercalation process. J. Appl. Polym. Sci. **1999**, *71* (7), 1133–1138.
17. Vaia, R.A.; Giannelis, E.P. Lattice of polymer melt intercalation in organically-modified layered silicates. Macromolecules **1997**, *30* (25), 7990–7999.
18. Vaia, R.A.; Giannelis, E.P. Polymer melt intercalation in organically-modified layered silicates: Model predictions and experiment. Macromolecules **1997**, *30* (25), 8000–8009.
19. Krishnamoorti, R.; Vaia, R.A.; Giannelis, E.P. Structure and dynamics of polymer-layered silicate nanocomposites. Chem. Mater. **1996**, *8* (8), 1728–1734.
20. Hackett, E.; Manias, E.; Giannelis, E.P. Molecular dynamics simulations of organically modified layered silicates. J. Chem. Phys. **1998**, *108* (17), 7410–7415.
21. Hackett, E.; Manias, E.; Giannelis, E.P. Computer simulation studies of PEO/layer silicate nanocomposites. Chem. Mater. **2000**, *12* (8), 2161–2167.
22. Lyatskaya, Y.; Balazs, A.C. Modeling the phase behavior of polymer–clay composites. Macromolecules **1998**, *31* (19), 6676–6680.
23. Carrado, K.A. Synthetic organo- and polymer-clays:

Preparation, characterization, and materials applications. Appl. Clay Sci. **2000**, *17* (1–2), 1–23.
24. Ishida, H.; Campbell, S.; Blackwell, J. General approach to nanocomposite preparation. Chem. Mater. **2000**, *12* (5), 1260–1267.
25. Zhao, B.; Brittain, W.J. Polymer brushes: Surface-immobilized macromolecules. Prog. Polym. Sci. **2000**, *25* (5), 677–710.
26. Kleinfeld, E.R.; Ferguson, G.S. Stepwise formation of multilayered nanostructural films from macromolecular precursors. Science **1994**, *265* (5170), 370–373.
27. Kotov, N.A.; Haraszti, T.; Turi, L.; Zavala, G.; Geer, R.E.; Dekany, I.; Fendler, J.H. Mechanism of and defect formation in the self-assembly of polymeric cation-montmorillonite ultrathin films. J. Am. Chem. Soc. **1997**, *119* (29), 6821–6832.
28. Fan, X.; Park, M.-K.; Xia, C.; Advincula, R.C. Surface structural characterization and mechanical testing by nanoindentation measurements of hybrid polymer/clay nanostructured multilayer films. J. Mater. Res. **2002**, *17* (7), 1622–1633.
29. Pinnavaia, T.J. Intercalated clay catalysts. Science **1983**, *220* (4595), 365–371.
30. Eastman, M.P.; Bain, E.; Porter, T.L.; Manygoats, K.; Whitehorse, R.; Parnell, R.A.; Hagerman, M.E. The formation of poly(methyl methacrylate) on transition metal-exchanged hectorite. Appl. Clay Sci. **1999**, *15* (1–2), 173–185.
31. Blumstein, A. Polymerization of adsorbed monolayers. I. Preparation of the clay–polymer complex. J. Polym. Sci. **1965**, *A3*, 2653–2664.
32. Solomon, D.H.; Rosser, M.J. Reactions catalyzed by minerals. Part I. Polymerization of styrene. J. Appl. Polym. Sci. **1965**, *9*, 1261–1271.
33. Lee, D.C.; Jang, L.W. Preparation and characterization of PMMA–clay hybrid composite by emulsion polymerization. J. Appl. Polym. Sci. **1996**, *61* (7), 1117–1122.
34. Noh, M.W.; Lee, D.C. Synthesis and characterization of PS–clay nanocomposite by emulsion polymerization. Polym. Bull. **1999**, *42* (5), 619–626.
35. Bandyopadhyay, S.; Giannelis, E.P.; Hsieh, A.J. Thermal and thermo-mechanical properties of PMMA nanocomposites. PMSE Prepr. **2000**, *82*, 208–209.
36. Al-Esaimi, M.M. Reaction catalyzed by montmorillonite: Polymerization of methyl methacrylate. J. Appl. Polym. Sci. **1997**, *64* (2), 367–372.
37. Tabtiang, A.; Lumlong, S.; Venables, R.A. The influence of preparation method upon the structure and relaxation characteristics of poly(methyl methacrylate)/clay composites. Eur. Polym. J. **2000**, *36* (12), 2559–2568.
38. Chen, G.; Ma, Y.; Qi, Z. Preparation and morphological study of an exfoliated polystyrene/montmorillonite nanocomposite. Scr. Mater. **2001**, *44* (1), 125–128.
39. Okamoto, M.; Morita, S.; Taguchi, H.; Kim, Y.H.; Kotaka, T.; Tateyama, H. Synthesis and structure of smectic clay/poly(methyl methacrylate) and clay/polystyrene nanocomposites via in situ intercalative polymerization. Polymer **2000**, *41* (10), 3887–3890.
40. Okamoto, M.; Morita, S.; Kim, Y.H.; Kotaka, T.; Tateyama, H. Dispersed structure change of smectic clay/poly(methyl methacrylate) nanocomposites by copolymerization with polar comonomers. Polymer **2001**, *42* (3), 1201–1206.
41. Dietsche, F.; Thomann, Y.; Thomann, R.; Mulhaupt, R. Translucent acrylic nanocomposites containing anisotropic laminated nanoparticles derived from intercalated layered silicates. J. Appl. Polym. Sci. **2000**, *75* (3), 396–405.
42. Doh, J.G.; Cho, I. Synthesis and properties of polystyrene–organoammonium montmorillonite hybrid. Polym. Bull. **1998**, *41* (5), 511–518.
43. Noh, M.H.; Jang, L.W.; Lee, D.C. Intercalation of styrene–acrylonitrile copolymer in layered silicate by emulsion polymerization. J. Appl. Polym. Sci. **1999**, *74* (1), 179–188.
44. Noh, M.H.; Lee, D.C. Comparision of characteristics of SAN–MMT nanocomposites prepared by emulsion and solution polymerization. J. Appl. Polym. Sci. **1999**, *74* (12), 2811–2819.
45. Biasci, L.; Agietto, M.; Ruggeri, G.; Ciardelli, R. Functionalization of montmorillonite by methyl methacrylate polymers containing side-chain ammonium cations. Polymer **1994**, *35* (15), 3296–3304.
46. Akelah, A.; Moet, A. Polymer–clay nanocomposites: Free-radical grafting of polystyrene on to organophilic montmorillonite interlayers. J. Mater. Sci. **1996**, *31* (13), 3589–3596.
47. Moet, A.; Akelah, A. Polymer–clay nanocomposites: Polystyrene grafted onto organophilic montmorillonite interlayers. Mater. Lett. **1993**, *18* (1–2), 97–102.
48. Dietsche, F.; Thomann, R.; Doell, W.; Mulhaupt, R. Acrylic nanocomposites containing organoclay nanofillers: Morphology development, properties and fracture behavior. PMSE Prepr. **2000**, *82*, 222–223.
49. Fu, X.; Qutubuddin, S. Synthesis of polystyrene–clay nanocomposites. Mater. Lett. **2000**, *42* (1–2), 12–15.
50. Fu, X.; Qutubuddin, S. Polymer–clay nanocomposites: Exfoliation of organophilic montmorillonite layers in polystyrene. Polymer **2001**, *42* (2), 807–813.
51. Zhu, J.; Morgan, A.B.; Lamelas, F.J.; Wilkie, C.A.

Fire properties of polystyrene–clay nanocomposites. Chem. Mater. **2001**, *13* (10), 3774–3780.
52. Yeh, J.-M.; Liou, S.-J.; Lin, C.-Y.; Cheng, C.-Y.; Chang, Y.-W.; Lee, K.-R. Anticorrosively enhanced PMMA–clay nanocomposite materials with quaternary alkylphosphonium salt as an intercalating agent. Chem. Mater. **2002**, *14* (1), 154–161.
53. Kim, Y.K.; Choi, Y.S.; Wang, K.H.; Chung, I.J. Synthesis of exfoliated PS/Na-MMT nanocomposites via emulsion polymerization. Chem. Mater. **2002**, *14* (12), 4990–4995.
54. Choi, Y.S.; Choi, M.H.; Wang, K.H.; Kim, S.O.; Kim, Y.K.; Chung, I.J. Synthesis of exfoliated PMMA/Na-MMT nanocomposites via soap-free emulsion polymerization. Macromolecules **2001**, *34* (26), 8978–8985.
55. Zeng, C.; Lee, L.J. Poly(methyl methacrylate) and polystyrene/clay nanocomposites prepared by in-situ polymerization. Macromolecules **2001**, *34* (12), 4098–4103.
56. Wang, D.; Zhu, J.; Yao, Q.; Wilkie, C.A. A comparison of various methods for the preparation of polystyrene and poly(methyl methacrylate) clay nanocomposites. Chem. Mater. **2002**, *14* (9), 3837–3843.
57. Muzny, C.D.; Butler, B.D.; Hanley, H.J.M.; Tsvetkov, F.; Peiffer, D.G. Clay platelet dispersion in a polymer matrix. Mater. Lett. **1996**, *28* (4–6), 379–384.
58. Choi, Y.S.; Wang, K.H.; Xu, M.; Chung, I.J. Synthesis of exfoliated polyacrylonitrile/Na-MMT nanocomposites via emulsion polymerization. Macromolecules **2002**, *14* (7), 2936–2939.
59. Hoffmann, B.; Dietrich, C.; Thomann, R.; Friedrich, C.; Mulhaupt, R. Morphology and rheology of polystyrene nanocomposites based upon organoclay. Macromol. Rapid Commun. **2000**, *21* (1), 57–61.
60. Zilg, C.; Thomann, R.; Baumert, M.; Finter, J.; Mulhaupt, R. Organic/inorganic hybrid materials and nanocomposites based upon layered silicates modified with cyclic amides. Macromol. Rapid Commun. **2000**, *21* (17), 1214–1219.
61. Zilg, C.; Dietsche, F.; Hoffmann, B.; Dietrich, C.; Mulhaupt, R. Nanofillers based upon organic layered silicates. Macromol. Symp. **2001**, *169* (1), 65–77.
62. Prucker, O.; Ruhe, J. Synthesis of poly(styrene) monolayers attached to high surface area silica gels through self-assembled monolayers of azo initiators. Macromolecules **1998**, *31* (3), 592–601.
63. Prucker, O.; Ruhe, J. Mechanisms of radical chain polymerizations initiated by azo compounds covalently bound to the surface of spherical particles. Macromolecules **1998**, *31* (3), 602–613.
64. Prucker, O.; Ruhe, J. Polymer layers through self-assembled monolayers of initiators. Langmuir **1998**, *14* (24), 6893–6898.
65. Huang, W.; Skanth, G.; Baker, G.L.; Bruening, M.L. Surface-initiated thermal radical polymerization on gold. Langmuir **2001**, *17* (5), 1731–1736.
66. Schmidt, R.; Zhao, T.; Green, J.-B.; Dyer, D.J. Photoinitiated polymerization of styrene from self-assembled monolayers on Gold. Langmuir **2002**, *18* (4), 1281–1287.
67. Dekking, H.G.G. Propagation of vinyl polymers on clay surfaces. I. Preparation, structure, and decomposition of clay initiators. J. Appl. Polym. Sci. **1965**, *9*, 1641–1651.
68. Dekking, H.G.G. Propagation of vinyl polymers on clay surfaces. II. Polymerization of monomers initiated by free radicals attached to clay. J. Appl. Polym. Sci. **1967**, *11*, 23–36.
69. Meier, L.; Shelden, R.; Caseri, W.; Suter, U. Polymerization of styrene with initiator ionically bound to high surface area mica: Grafting via an unexpected mechanism. Macromolecules **1994**, *27* (6), 1637–1642.
70. Huang, X.; Brittain, W.J. Synthesis and characterization of PMMA nanocomposites by suspension and emulsion polymerization. Macromolecules **2001**, *34* (10), 3255–3260.
71. Fan, X.; Xia, C.; Fulghum, T.; Park, M.-K.; Locklin, J.; Advincula, R.C. Polymer layer grafted from clay surface on planar substrate through free radical surface initiated polymerization (SIP). Langmuir **2003**, *19* (3), 916–923.
72. Zhou, Q.; Wang, S.; Fan, X.; Advincula, R.C.; Mays, J. Living anionic surface-initiated polymerization (LASIP) of a polymer on silica nanoparticles. Langmuir **2002**, *18* (8), 3324–3331.
73. Fan, X.; Xia, C.; Advincula, R.C. Grafting of polymers from clay nanoparticles via in situ free radical surface-initiated polymerization: Monocationic versus bicationic initiators. Langmuir **2003**, *19* (10), 4381–4389.
74. Fan, X.; Xia, C.; Advincula, R.C. Poly(methyl methacrylate) (PMMA)–clay nanocomposites by bulk, solution, and suspension surface initiated polymerization (SIP). Polym. Prepr. **2003**, *44* (1), 1099–1100.
75. Fan, X.; Xia, C.; Advincula, R.C. Poly(methyl methacrylate) (PMMA)–clay nanocomposites by bulk, solution, and suspension surface initiated polymerization. (SIP). Macromolecules. submitted.
76. Zhou, Q.; Wang, S.; Fan, X.; Advincula, R.C.; Mays, J. Living anionic surface-initiated polymerization (LASIP) of a polymer on silica nanoparticles. Langmuir **2002**, *18* (8), 3324–3331.
77. Zhou, Q.; Fan, X.; Xia, C.; Mays, J.; Advincula,

R.C. Living anionic surface initiated polymerization (SIP) of styrene from clay surfaces. Chem. Mater. **2001**, *13* (8), 2465–2467.
78. Fan, X.; Zhou, Q.; Xia, C.; Crsitopholi, W.; Mays, J.; Advincula, R.C. Living anionic surface-initiated polymerization (LASIP) of styrene from clay nanoparticles using surface bound 1,1-diphenylethylene (DPE) initiators. Langmuir **2002**, *18* (11), 4511–4518.
79. Weimer, M.W.; Chen, H.; Giannelis, E.P.; Sogah, D.Y. Direct synthesis of dispersed nanocomposites by in situ living free radial polymerization using a silicate-anchored initiator. J. Am. Chem. Soc. **1999**, *121* (7), 1615–1616.
80. Usuki, A.; Kojima, Y.; Kawasumi, M.; Okada, A.; Kurauchi, T.; Kamigaito, O. Characterization and properties of nylon 6-clay hybrid. Polym. Prepr. **1990**, *31*, 651–652.
81. Usuki, A.; Kawasumi, M.; Kojima, Y.; Okada, A.; Kurauchi, T.; Kamigaito, O. Swelling behavior of montmorillonite cation exchanged for ω-amino acids by ϵ-aprolactam. J. Mater. Res. **1993**, *8* (5), 1174–1178.
82. Okada, A.; Usuki, A. The chemistry of polymer–clay hybrids. Mater. Sci. Eng. **1995**, *C3* (1), 109–115.
83. Messersmith, P.B.; Giannelis, E.P. Synthesis and barrier properties of poly(ϵ-caprolactone)-layered silicate nanocomposites. J. Polym. Sci., A, Polym. Chem. **1995**, *33* (7), 1047–1057.
84. Messersmith, P.B.; Giannelis, E.P. Polymer-layered silicate nanocomposites: In situ intercalative polymerization of ϵ-caprolactone in layered silicates. Chem. Mater. **1993**, *5* (8), 1064–1066.
85. Krishnamoorti, R.; Giannelis, E.P. Rheology of end-tethered polymer layered silicate nanocomposites. Macromolecules **1997**, *30* (14), 4097–4102.
86. Krishnamoorti, R.; Yurekli, K. Rheology of polymer layered silicate nanocomposites. Curr. Opin. Colloid Interface Sci. **2001**, *6*, 464–470.
87. Lofgren, A.; Albertsson, A.-C.; Dubois, P.; Jerome, R. Recent advances in ring-opening polymerization of lactones and related compounds. J. Macromol. Sci., Rev. Macromol. Chem. Phys. **1995**, *C35*, 379–418.
88. Mecerreyes, D.; Jerome, R.; Dubois, P. Novel macromolecular architectures based on aliphatic polyesters: Relevance of the coordination-insertion ring opening polymerization. Adv. Polym. Sci. **1999**, *147*, 1–59.
89. Lepoittevin, B.; Pantoustier, N.; Devalckenaere, M.; Alexandre, M.; Kubies, D.; Calberg, C.; Jerome, R.; Dubois, P. Poly(ε-caprolactone)/clay nanocomposites by in-situ intercalative polymerization catalyzed by dibutyltin dimethoxide. Macromolecules **2002**, *35* (22), 8385–8390.
90. Lan, T.; Kaviratna, P.D.; Pinnavaia, T.J. Mechanism of clay tactoid exfoliation in epoxy–clay nanocomposites. Chem. Mater. **1995**, *7* (12), 2144–2150.
91. Messersmith, P.B.; Giannelis, E.P. Synthesis and characterization of layered silicate–epoxy nanocomposites. Chem. Mater. **1994**, *6* (10), 1719–1725.
92. Kornmann, X.; Lindberg, H.; Berglund, L.A. Synthesis of epoxy–clay nanocomposites: Influence of the nature of the clay on structure. Polymer **2001**, *42* (4), 1303–1310.
93. Kornmann, X.; Berglund, L.A.; Sterte, J.; Giannelis, E.P. Nanocomposites based on montmorillonite and unsaturated polyester. Polym. Eng. Sci. **1998**, *38* (8), 1351–1358.
94. Wang, Z.; Pinnavaia, T.J. Nanolayer reinforcement of elastomeric polyurethane. Chem. Mater. **1998**, *10* (12), 3769–3771.
95. Wang, Z.; Pinnavaia, T.J. Hybrid organic–inorganic nanocomposites: Exfoliation of magadiite nanolayers in an elastomeric epoxy polymer. Chem. Mater. **1998**, *10* (7), 1820–1826.
96. Alexandre, M.; Dubois, P.; Sun, T.; Garces, J.M.; Jerome, R. Polyethylene-layered silicate nanocomposites prepared by the polymerization-filling technique: Synthesis and mechanical properties. Polymer **2002**, *43* (7), 2123–2132.
97. Sun, T.; Garces, J.M. High-performance polypropylene–clay nanocomposites by in-situ polymerization with metallocene/clay catalysts. Adv. Mater. **2002**, *14* (2), 128–130.
98. Hong, S.H.; Kim, B.H.; Joo, J.; Kim, J.W.; Choi, H.J. Polypyrrole–montmorillonite nanocomposites synthesized by emulsion polymerization. Curr. Appl. Phys. **2001**, *1* (6), 447–450.
99. Kim, B.H.; Jung, J.H.; Kim, J.W.; Choi, H.J.; Joo, J. Physical characterization of polyaniline-Na^+-montmorillonite nanocomposite intercalated by emulsion polymerization. Synth. Met. **2001**, *117* (1–3), 115–118.
100. Lee, D.; Char, K. Thermal degradation behavior of polyaniline in polyaniline/Na^+-montmorillonite nanocomposites. Polym. Degrad. Stab. **2002**, *75* (3), 555–560.

Polymeric and Biomolecular Nanostructures: Fabrication by Scanning Probe Lithography

Stefan Zauscher
Duke University, Durham, North Carolina, U.S.A.

INTRODUCTION

One central goal of materials engineering is to produce hierarchical materials that are ordered over a range of length scales and in which larger-scale structural and physicochemical properties are controlled by molecular characteristics.[1] Fabricating molecular, polymeric, and biomolecular structures on surfaces and controlling their architecture on the nanometer length scale is important for a wide range of applications. For example, the ability to pattern surfaces with polymers and biomacromolecules has important applications in biosensors,[2] proteomic chips,[3] and nanofluidic devices.[4] To date, fabrication of micropatterns and microstructures for these applications has relied on either photolithography[2,5,6] or an ensemble of related techniques termed "soft lithography"[7] that use elastomeric stamps or molds. These techniques have been extremely successful at the micrometer length scale or larger, but have a number of inherent limitations that preclude their extension to the nanometer length scale. X-ray lithography[8–10] requires expensive mask alignment equipment, and the photoresists and developers commonly used are incompatible with many organic and biomolecules. Furthermore, like electron beam lithography,[8,11,12] it requires expensive ultrahigh vacuum equipment.

This review will focus on methods that allow for molecular-level control in the fabrication of polymeric and biomolecular nanostructures using scanning probe lithography (SPL,[13–15]). Among all the available nanofabrication techniques, SPL has unique advantages because of its simplicity and capability to both manipulate and image nanostructures on surfaces. In contrast to other nanofabrication methods, SPL, which is capable of resolution less than 50 nm[16] is technologically simpler, significantly cheaper, and therefore a much more widely accessible method to most researchers, requiring only a scanning probe microscope (SPM) such as a scanning tunneling microscope (STM.[17]), an atomic force microscope (AFM[18]), or a near-field scanning optical microscope (NSOM[19,20]) (Fig. 1).

SPL can behave like an interactive plotter, registering with nanostructures on a surface, while offering resolution comparable to electron beam lithography[21] The slow-speed serial nature of the lithography process can be circumvented by employing cantilever arrays.[22]

METHODOLOGY

Direct manipulation of both self-assembly and supramolecular chemistry are critical to materials engineering on the nanoscale. Generally, two fabrication approaches are used: a "bottom-up" or "grafting from" approach, which constructs supramolecular ensembles by synthesis of functionalized, small molecular units, using, for example, surface initiated polymerization; and a "grafting to" approach that seeks to directly manipulate and position macromolecules, proteins, or ligands on a surface with nanometer precision.

Self-Assembled Affinity Templates

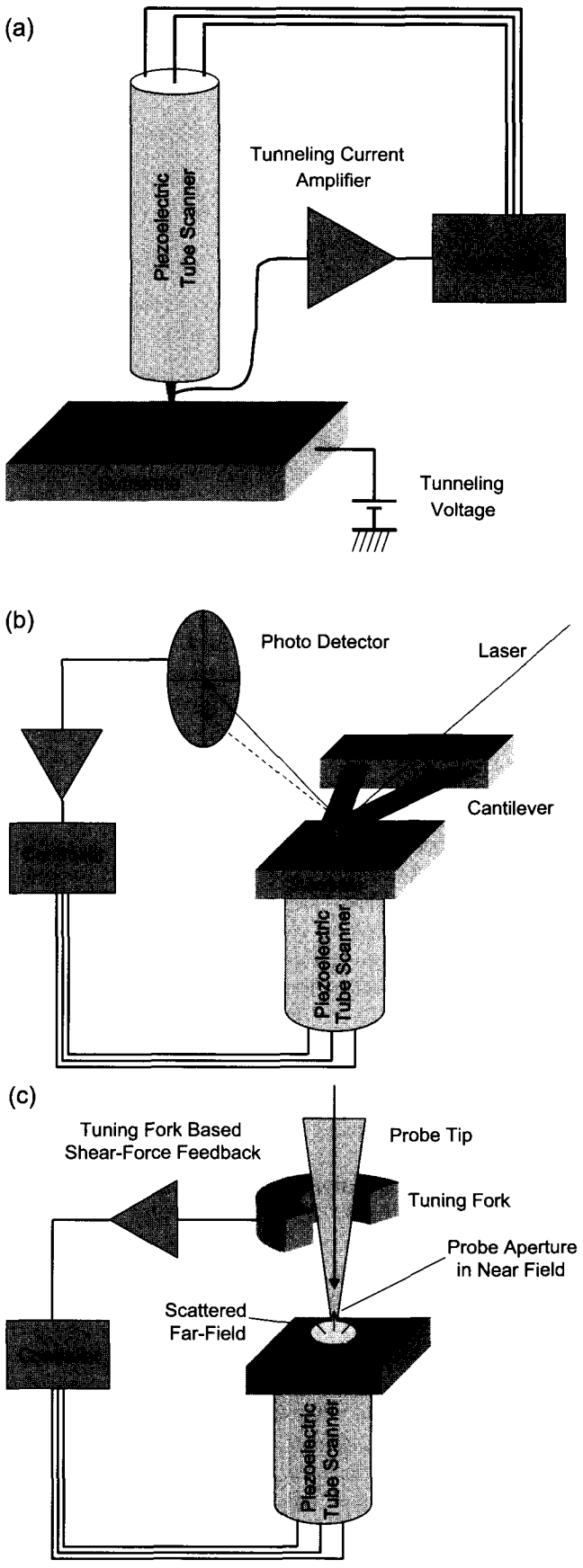
Highly ordered and spatially confined self-assembled monolayers (SAMs) serve as initiator platforms for polymerization of polymer brushes and as growth templates for directed self-assembly of synthetic and biomolecules. The molecules forming these "affinity templates" are typically bifunctional, where the surface binding group (silane[23] or thiol[24,25]) is chosen to match the substrate chemistry. The linker, usually an inert hydrocarbon chain, connects the surface binding group with the ligating, functional head group, often a hydroxyl (–OH) or carboxyl (–COOH) group.

Surface-Initiated Polymerization

Attaching polymer brushes[26] with thicknesses on the molecular scale to solid surfaces allows one to tailor the surface properties of materials by imparting desirable energetic, mechanical, and electrical functionalities. The in situ formation of polymer brushes in a "grafting-from" approach, using patterned SAMs of an initiator, provides spatial flexibility and achieves higher packing densities than possible in a "grafting-to" approach.[27,28]

Seminal work by Prucker and Rühe[27,28] led the way for generating polymer brushes in a "grafting from"

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120014195



approach using free radical polymerization. Although the templated fabrication of polymeric nanostructures from the "bottom-up" has been prototypically demonstrated, and many methods to initiate polymerization reactions have been used (e.g., anion,[29] cation,[30] radical,[27] plasma,[31] condensation,[32] photochemical,[33] electrochemical,[34] and ring-opening metathesis polymerization.[35–37]), preparation of precisely patterned, surface-attached polymeric structures with controlled lengths, conformational geometries, functionality, and properties is still in its infancy. So far, atom-transfer radical polymerization (ATRP) has been the workhorse polymerization methodology used by researchers attempting to prepare surface-attached polymer brushes of controlled structure[38,39] This transition metal-based, controlled radical polymerization chemistry produces functional polymers with defined molecular weight and polydispersity, and as a result of the "living" nature of the catalyst, allows the ready synthesis of block copolymers.

Protein Grafting

The precision with which one can organize large molecular structures, such as proteins, depends on the quality and chemical functionality of the affinity template nanostructure.[40] Ideally, the template surface and background polymers are chosen so that the protein to be grafted selectively links either with the template or the background only. Some approaches for protein grafting or "immobilization" rely on physisorption, while others make use of chemical bonds between protein and substrate, particularly if the patterned protein is subjected to hydrodynamic or other physical forces. Covalent immobilization methods include formation of amide, imine, and disulfide bonds. While these approaches are relatively simple, they often preclude direct control over protein orientation on the surface, which is problematic if protein binding to its recognition partners is desired. Recently, Hyun et al.[41] developed a new nanofabrication method allowing control over molecular orientation by using molecular recognition mediated protein immobilization.

SCANNING PROBE LITHOGRAPHY

Shortly after the invention of the scanning probe microscope, it was realized that the tips used for imaging could also be used to modify the surfaces. Since then, many SPL techniques have been developed based on various

Fig. 1 Schematic view of scanning probe microscopes. (a) Scanning tunneling microscope (STM), (b) atomic force microscope (AFM), and (c) near-field scanning optical microscope (NSOM). (*View this art in color at www.dekker.com.*)

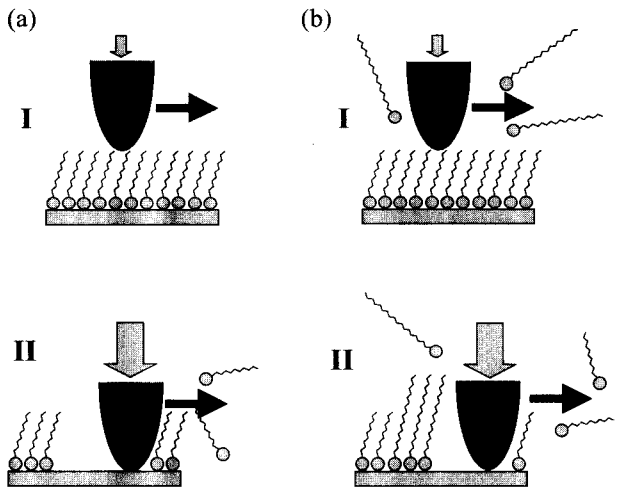


Fig. 2 Schematic diagrams of two basic manipulation mechanisms using AFM: (a) nanoshaving and (b) nanografting. The corresponding imaging (I) and fabrication (II) modes are depicted in the top and bottom rows, respectively. (From Ref. [44]. Copyright 2000, American Chemical Society.) (*View this art in color at www.dekker.com.*)

chemical, physical, and electrical modifications of surfaces, including mechanical scratching,[42–44] electrochemical anodization of silicon surfaces,[45,46] decomposition of self-assembled monolayers,[45,47–49] electric field induced chemical reactions,[50,51] and electrochemical reactions in solution using electrochemical STM tips.[52–57] Several comprehensive reviews of SPM related lithography can be found in the literature.[16,58]

Depending on the instrument design, either the probe tip scans relative to a stationary sample surface, or the surface is scanned relative to a stationary tip. Scanning is typically achieved through piezoelectric actuators. The requirement for accurate tip positioning has driven SPM design to use closed-loop scanning to compensate for the nonlinear and often hysteretic behavior of piezoelectric actuators. Increasing interest in scanning probe lithography has also spurred the development of user interfaces that allow the user to program arbitrary motions of the tip relative to the sample,[59] and several SPM manufacturers now offer lithography packages. Current methods for fabricating polymer nanostructures include nanoshaving and nanografting, dip-pen nanolithography (DPN), direct-write DPN, and near-field scanning optical lithography.

Nanoshaving and Nanografting

Perhaps the simplest method for fabricating patterned substrates is nanoshaving, in which SAM adsorbates are mechanically displaced while the cantilever tip is dragged over the surface, leaving behind a patterned surface.[42,44,60] Large contact pressure during patterning[44] creates large shear forces that exceed the displacement threshold of the usually chemisorbed molecules (Fig. 2a). In nanografting, the AFM tips are again used to displace thiol molecules from adsorptions sites; however, cantilever and SAM are now immersed in a solution of a different thiol[43] (Fig. 2b).

These thiols in solution immediately self-assemble on the newly exposed gold surface left behind as the tip shaves the matrix SAM. The resulting SAM patterns can now be used as affinity templates for protein immobilization[40] or for surface initiated nanopolymerization.

Dip-Pen Nanolithography

More recently, a new SPL technique, "dip-pen" nanolithography (DPN),[61,62] has permitted patterning of monolayer films of thiolated organic molecules with sub-100-nm resolution onto gold substrates. The technology uses the spontaneous condensation of water vapor between the AFM tip and substrate to transport organic molecules from the AFM tip to the surface in the area specifically defined by the tip–surface interaction[63] (Fig. 3).

To draw on a familiar analogy, the AFM tip acts as a "pen," the organic molecules act as "ink," and the surface acts as "paper" to "draw" the nanostructures. DPN allows direct patterning of nanoscale structures in both serial and parallel fashions.[64] Attractive features of DPN over most other SPL techniques are its simplicity—it requires no liquid cells—and its ability to create various nanostructures in a "direct-write" fashion. Furthermore, as long as the tip can be inked with a solution containing the reactant of interest and a capillary meniscus main-

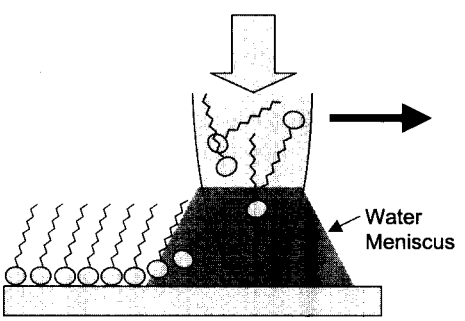


Fig. 3 Schematic representation of dip-pen nanolithography (DPN). A water meniscus forms between the AFM tip, typically coated with a ω-functionalized alkane thiol (e.g., mercaptohexadecanoic acid) and a gold substrate. The size of the meniscus, which is controlled by relative humidity, affects the thiol transport rate, the effective tip–substrate contact area, and DPN resolution. (*View this art in color at www.dekker.com.*)

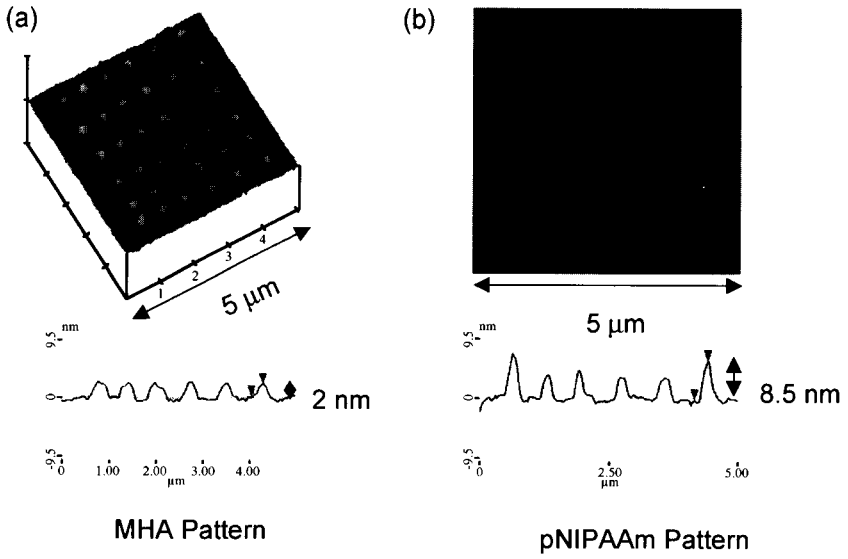


Fig. 4 Surface initiated nanopolymerization (SINP) of *N*-isopropylacrylamide (NIPAAm) on a mercaptohexadecanoic acid (MHA) pattern, conjugated with free radical initiator. (a) AFM lateral force height image of MHA patterned with DPN and typical cross section (contact mode in air). (b) AFM height image of the corresponding pNIPAAm pattern and typical cross section (TappingMode™ in air). The feature height after 30 hr reaction time ranges between 7 and 9 nm, and is amplified at the pattern boundary. (From Ref. [65]. Copyright 2003, Materials Research Society.) (*View this art in color at www.dekker.com.*)

tained between the tip and surface, there appears to be little limitation on the types of molecules that can be transferred onto a surface. Thus DPN enables a large number of additive chemical reactions to be performed on diverse substrates with nanometer resolution comparable to that achieved with electron-beam lithography. However, to achieve this nanometer resolution, DPN requires molecularly smooth gold substrates.

Recently, Ahn et al.[65] combined surface initiated polymerization with DPN in a fabrication process termed "surface initiated nanopolymerization" (SINP) to fabricate patterned, surface-confined, stimulus-responsive polymer brushes (Fig. 4). In this approach, mercaptohexadecanoic acid (MHA) was patterned on gold thin films by DPN (Fig. 4a) and subsequently conjugated with a free radical initiator. The cross-sectional profile of a pattern with 300-nm dots reveals that the polymer brush height at the periphery of the pattern is larger than that in the center (Fig. 4b).

These and other results suggest that the rate of polymerization, and thus the brush height in uncontrolled free radical polymerizations of *N*-isopropylacrylamide (NIPAAm), are affected by pattern size. This may likely be attributable to the density of initiator immobilized on the surface and limited monomer diffusion to the surface of the growing brush (1-D).[28] Diffusional freedom is larger at pattern boundaries (2-D) and polymer growth is usually larger there.

Controlled/living radical polymerization methods, such as atom transfer radical polymerization (ATRP), decrease the probability of bimolecular termination reactions, yielding polymer brushes of uniform molecular weight and controlled composition. Kaholek et al.[66] succeeded in fabricating stimulus-responsive polymer patterns with linewidths of 200 nm, by directly patterning the initiator thiol using DPN, and subsequent amplification of the pattern using ATRP of NIPAAm in an organometallic catalyst solution (Fig. 5).

The preparation of patterned protein surfaces using DPN was first demonstrated by Demers et al.[67] and Lee et al.[68] who physisorbed lysozyme and immunoglobulin G (IgG) onto DPN nanopatterned mercaptohexadecanoic acid (MHA) templates, achieving feature sizes on the order of 100 nm. By backfilling with a protein-resistant oligoethylene glycol-terminated alkanethiol SAM,[69] they avoided nonspecific protein adsorption that could have led to large background signals and complicated differentiation between active and inactive areas of the substrate.

A method for molecular-recognition mediated, stepwise fabrication of patterned protein nanostructures with feature sizes on the order of 200 nm was recently developed in the author's laboratory (Fig. 6).[41]

In this method, an amine-terminated biotin derivative was covalently conjugated with a chemically activated MHA SAM nanopattern, patterned with DPN on gold. The patterned surface was then incubated with streptavidin whose binding to the pattern was mediated by molecular recognition between biotin and streptavidin. The patterned streptavidin surface functioned as a universal

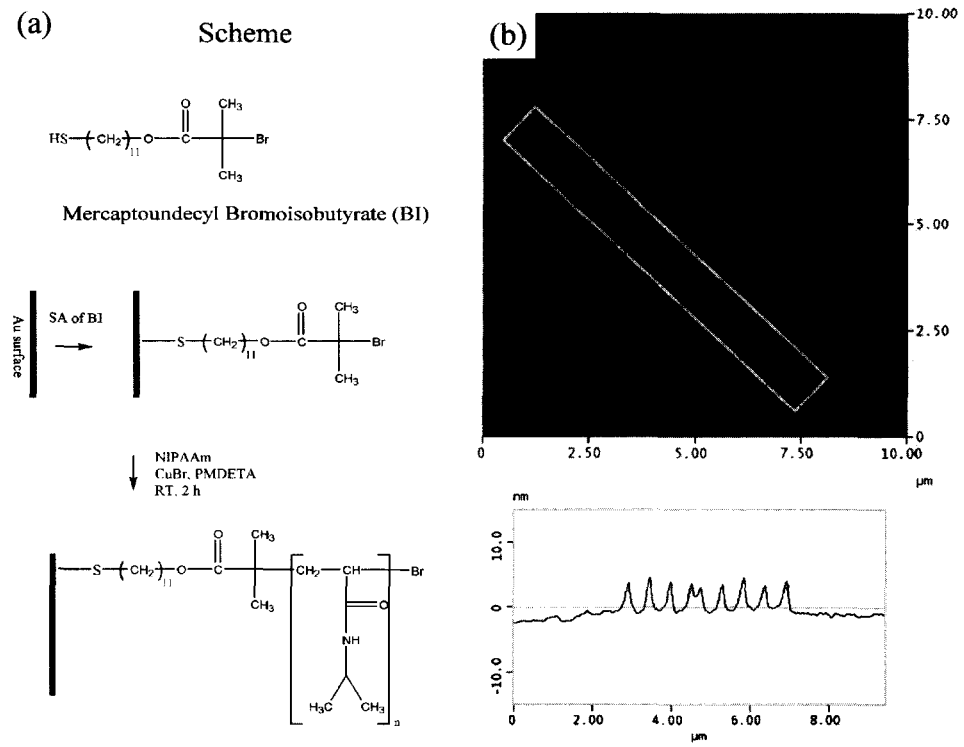


Fig. 5 Surface initiated nanopolymerization of NIPAAm using atom transfer radical polymerization (ATRP) on directly patterned initiator thiol. (a) Reaction scheme. (b) AFM height image of the resulting pNIPAAm line pattern showing a line width of 220 nm line and a brush height of 5.7 nm (TappingMode™ in Milli-Q™ water). (*View this art in color at www.dekker.com.*)

platform for molecular-recognition mediated protein immobilization through use of biotinylated molecules. "Smart," reversible protein nanopatterns were fabricated by using iminobiotin as the molecular recognition partner. These structures are pH-sensitive and allow for the reversible immobilization of iminobiotin tagged proteins. This latter characteristic makes possible the regeneration of the nanopattern.

The concept of "smart," regenerable protein patterns was carried one step further by Hyun et al.,[70] who developed a new method for the fabrication of stimulus-responsive polypeptide nanostructures using DPN. They

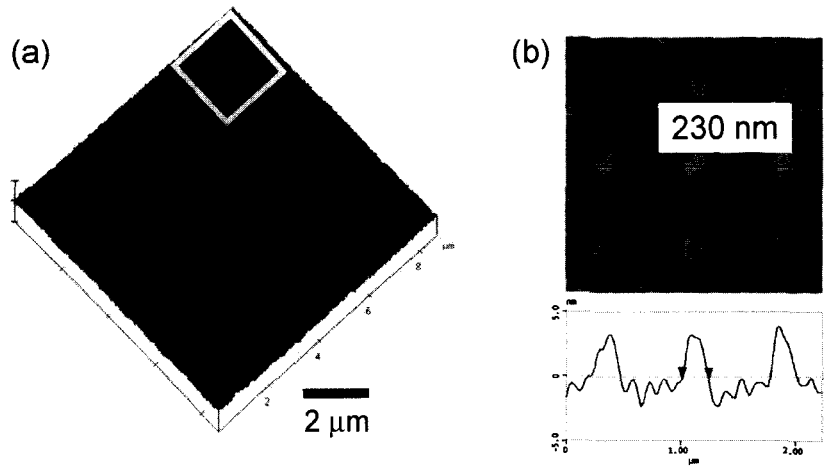


Fig. 6 AFM height images of Biotin–BSA nanopatterns (TappingMode™ in Milli-Q™ water). (a) 144 dot pattern with average feature size of 230 nm. (b) Zoom of the area indicated by the square in panel (a) and representative cross section showing the typical feature size. (From Ref. [41]. Copyright 2002, American Chemical Society.) (*View this art in color at www.dekker.com.*)

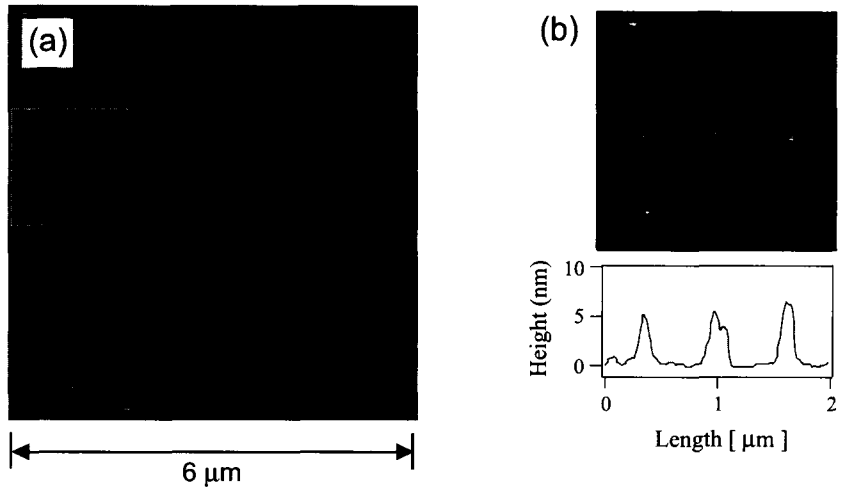


Fig. 7 (a) AFM TappingMode™ height image of a 10×9 elastin-like polypeptide (ELP) dot array in PBS buffer. (b) Enlarged view of area indicated in panel (a) and representative cross section, showing a feature size of about 200 nm. (*View this art in color at www.dekker.com.*)

created an array comprising elastin-like polypeptide (ELP) nanodots with a diameter of approximately 200 nm on gold against a nonfouling oligoethylene glycol-terminated alkanethiol SAM (Fig. 7).

ELPs, chemically grafted onto self-assembled monolayers on gold, exhibit an interfacial hydrophilic–hydrophobic phase transition in response to increased temperature or ionic strength. This stimulus-response behavior of ELP nanopatterns can be harnessed for the reversible immobilization of small amounts of ELP-tagged proteins from solution, which is important for protein detection and purification on the nanoscale.

Direct-Write DPN

Direct-write DPN of polymeric and biomolecular structures has great appeal as an efficient nanofabrication method. For example, tip-directed in-situ polymerization potentially eliminates multiple-step processing that currently requires amplification of a pattern in a bulk polymerization step. Even greater advantages present themselves for direct-write DPN of proteins in areas of drug discovery, diagnostics, and proteomics, as it offers the possibility of generating arrays made from more than one protein.

Until recently, DPN has been used to fabricate affinity arrays from small organic molecules that aid in the directed deposition of proteins and biomacromolecules from solution.[41,68,70] Demers et al.[67] developed a variation of DPN that directly writes oligonucleotides on metals and insulators. A key strategy for the success of this approach was the use of matched ink–substrate combinations, directing the transfer of oligonucleotides onto the substrate. Lee et al.[71] recently developed an innovative approach for directly writing proteins using DPN. This method hinges on the realization that protein transfer from the tip onto the substrate requires a biocompatible and hydrophilic (nonfouling) tip surface layer (Fig. 8).

The latter was achieved by tip-surface treatment with a poly(ethylene-glycol)silane. Directed, and orientationally

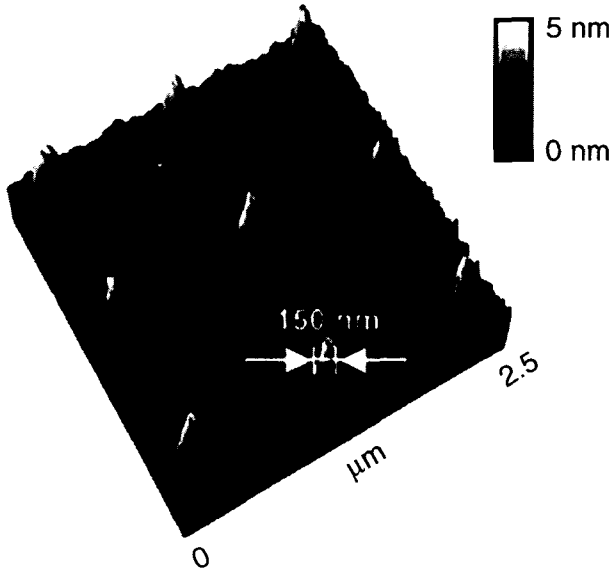


Fig. 8 Protein nanoarray prepared via direct-write DPN. Contact mode height image of a Lysozyme nanodot array with a feature size of about 150 nm. (From Ref. [71]. Copyright 2003 American Chemical Society.) (*View this art in color at www.dekker.com.*)

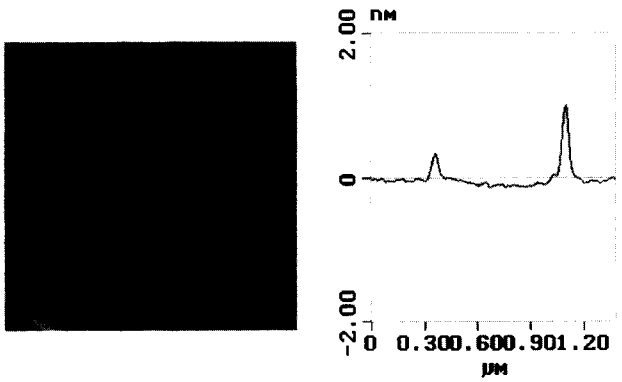


Fig. 9 Tip-directed electrochemical polymerization of 3, 4-ethylenedioxythiophene. AFM height image (contact mode in air) showing two polymer lines each about 50 nm wide. (From Ref. [74]. Copyright 2001, American Chemical Society.) (*View this art in color at www.dekker.com.*)

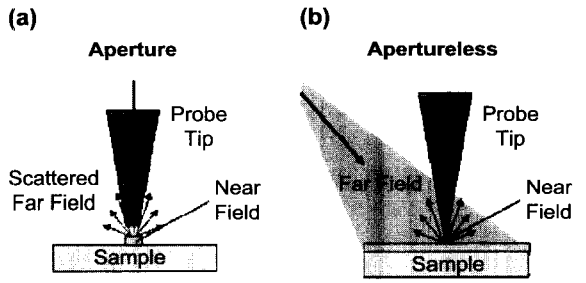


Fig. 10 Two possible configurations for scanning near-field optical lithography. (a) NSOM aperture probe where radiation enters the region of interest in the near field through the probe. (b) Aperture-less NSOM using far field illumination. Intensity amplification occurs beneath the probe tip in the optical near field. (*View this art in color at www.dekker.com.*)

controlled protein deposition on the substrate surface is possible by judicious matching of surface and protein functionality.[71,72]

Current photolithography, microcontact printing, template synthesis, and scanning electrochemical microlithography techniques can yield conducting polymer microstructures, but these techniques possess significant limitations for patterning structures of <100-nm dimensions. The availability of a wide range of monomers enables direct-writing of polymeric materials with a wide range of interesting electronic and electrooptical properties at the nanoscale level. Li et al.[73] have recently introduced a variation of DPN, in which the water meniscus on the AFM tip is not only used as the transfer medium but also serves as a nanometer-sized electrochemical cell in which metal salts can be dissolved, electrochemically reduced, and ultimately deposited on a surface.[73] They showed that electrochemical DPN can be used to directly fabricate conducting polythiophene nanostructures on semiconducting and insulating surfaces in the sub-100-nm size scale.[74] To pattern these nanostructures, a negative bias voltage was applied between the AFM tip and a silicon surface, and the tip was translated across the surface in a preprogrammed pattern. The applied voltage electrochemically polymerized the monomer, resulting in tip-defined deposition of poly-3,4-ethylenedioxythiophene on the substrate (Fig. 9).

Near-Field Scanning Optical Lithography

To date, photolithography is perhaps the most successful microfabrication technology, as most microelectronic circuits are produced via photolithographic methods. In conventional photolithography, the lateral resolution is determined by the Abbe diffraction limit (a far-field effect), which limits the feature size to about half the wavelength of the radiation used. The diffraction limit can be circumvented if the photosensitive sample is placed in the optical near-field of the focusing elements.[19,75] Optical lithography in the evanescent near field is mainly developed as a static patterning process by using a shadow mask in close proximity to a photoresist layer during exposure.[76]

In near-field scanning optical lithography (NSOL), this approach is taken one step further by inducing photopolymerization only in the vicinity of the probe tip (Fig. 10a).[19,75] In most NSOL techniques, laser radiation is coupled directly into a tapered single-mode optical fiber where the small fiber aperture limits the intensity of the deliverable radiation. Another NSOL technique capitalizes on the field enhancement of optical radiation in the near field of a conducting SPM tip due to photon absorption (Fig. 10b).[77–79]

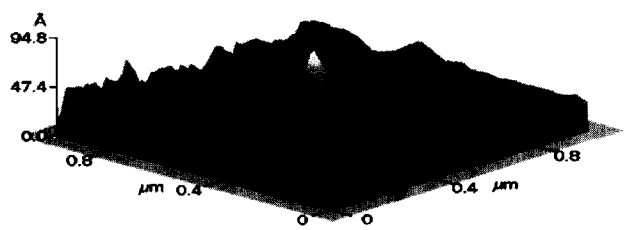


Fig. 11 Aperture-less near-field scanning optical lithography where the optical field enhancement at the tip initiates locally the photopolymerization process. AFM height image (Tapping-Mode™ in air) showing a nanometer scale polymer feature polymerized from a photo-polymerizable formulation (resist). (From Ref. [39]. Copyright 2000, Japanese Institute of Pure and Applied Physics.)

This approach is of considerable interest for surface nanopatterning and nanofabrication as it circumvents the use of delicate optical fibers and instead makes use of readily available laser sources and focusing optics.

Wurtz et al.[79] report on photopolymerization of a photoresist formulation induced by optical field enhancement in the near field of a conducting tip.[78,80] The free radical polymerization reaction was triggered by an actinic light activated sensitizer dye that, in the triplet state, reacts with amine to form free radicals. The threshold energy necessary to trigger a reaction is high and requires relatively large field intensities. Sufficiently large fields were only achieved in close proximity to the tip apex of a conducting AFM cantilever. Polymeric features with 70-nm lateral feature size were fabricated (Fig. 11).

These results are significant because they demonstrate that direct nanopolymerization in the evanescent near-field of a conducting AFM tip is possible, and affords an elegant way for patterning surfaces with polymers.

CONCLUSION

Advances in the area of scanning probe lithography have made available techniques, such as dip-pen nanolithography and nanografting, that allow fabrication of polymeric structures on surfaces on the nanoscale. While significant progress has been made in nanopatterning of biomacromolecules using affinity templates and direct writing, nanofabrication of polymeric structures from the "bottom up" using scanning probe lithography has only been prototypically demonstrated. Thus there is now a need to build more complex polymeric structures with precisely controlled lengths, conformational geometries, and chemical functionalities. In pursuit of these goals, techniques must be further developed which allow for the controlled synthesis of surface-attached polymers.

Several factors need to be addressed for the viability of SPL as a fabrication technology for polymeric nanostructures beyond the laboratory scale: 1) lateral resolution; 2) alignment; 3) reliability; and 4) throughput. While items 1)–3) are intrinsic capabilities associated with SPMs, a considerable restriction is presented by the slow, serial nature of the patterning process—one tip creates a single feature at a time. Recent developments of massively parallel proximal probe systems have been demonstrated,[22] and offer prospects toward achieving cost-effective, high-throughput "direct write" fabrication. The development of complex patterns is only possible through sufficient automation of the instrumentation and through pattern generation making use of computer-aided design (CAD). Progress in this area is of primary importance for the evolution of existing SPL nanofabrication strategies.

ACKNOWLEDGMENTS

The author would like to thank the National Science Foundation for support through grants NSF EEC-021059 and NSF DMR-0239769 CAREER AWARD. The author is grateful to Ken C. Caster for helpful discussions. The help of Chris Markley in preparing this manuscript is acknowledged.

REFERENCES

1. Dan, N. Synthesis of hierarchical materials. Trends Biotechnol. **2000**, *18* (9), 370–374.
2. Blawas, A.S.; Reichert, W.M. Protein patterning. Biomaterials **1998**, *19* (7–9), 595–609.
3. Mrksich, M.; Whitesides, G.M. Patterning self-assembled monolayers using microcontact printing—A new technology for biosensors. Trends Biotechnol. **1995**, *13* (6), 228–235.
4. Beebe, D.J.; Moore, J.S.; Yu, Q.; Liu, R.H.; Kraft, M.L.; Jo, B.H.; Devadoss, C. Microfluidic tectonics: A comprehensive construction platform for microfluidic systems. Proc. Natl. Acad. Sci. U. S. A. **2000**, *97* (25), 13488–13493.
5. Hengsakul, M.; Cass, A.E.G. Protein patterning with a photoactivatable derivative of biotin. Bioconjug. Chem. **1996**, *7* (2), 249–254.
6. Mooney, J.F.; Hunt, A.J.; McIntosh, J.R.; Liberko, C.A.; Walba, D.M.; Rogers, C.T. Patterning of functional antibodies and other proteins by photolithography of silane monolayers. Proc. Natl. Acad. Sci. U. S. A. **1996**, *93* (22), 12287–12291.
7. Xia, Y.N.; Whitesides, G.M. Soft lithography. Angew. Chem. Int. Ed. **1998**, *37* (5), 551–575.
8. Chen, Z.W. Nanometric-scale electron-beam lithography. Adv. Electron. El. Phys. **1992**, *83*, 107–201.
9. Tandon, U.S.; Pant, B.D.; Kumar. A. An overview of x-ray-lithography for use in semiconductor-device preparation. Vacuum **1991**, *42* (18), 1219–1228.
10. Tolfree, D.W.L. Microfabrication using synchrotron radiation. Rep. Prog. Phys. **1998**, *61* (4), 313–351.
11. Craighead, H.G.; Mankiewich, P.M. Ultra-small metal-particle arrays produced by high-resolution electron-beam lithography. J. Appl. Phys. **1982**, *53* (11), 7186–7188.
12. Pease, R.F.W. Nanolithography and its prospects as a manufacturing technology. J. Vac. Sci. Technol., B **1992**, *10* (1), 278–285.
13. Kim, Y.; Lieber, C.M. Machining oxide thin-films with an atomic force microscope—Pattern and

object formation on the nanometer scale. Science **1992**, *257* (5068), 375–377.
14. Wada, Y. Possible application of micromachine technology for nanometer lithography. Microelectron. J. **1998**, *29* (9), 601–611.
15. Quate, C.F. Scanning probes as a lithography tool for nanostructures. Surf. Sci. **1997**, *386* (1–3), 259–264.
16. Nyffenegger, R.M.; Penner, R.M. Nanometer-scale surface modification using the scanning probe microscope: Progress since 1991. Chem. Rev. **1997**, *97* (4), 1195–1230.
17. Binnig, G.; Rohrer, H. Scanning tunneling microscopy. Helv. Phys. Acta **1982**, *55* (6), 726–735.
18. Binnig, G.; Quate, C.F.; Gerber, C. Atomic force microscope. Phys. Rev. Lett. **1986**, *56* (9), 930–933.
19. Betzig, E.; Trautman, J.K. Near-field optics—Microscopy, spectroscopy, and surface modification beyond the diffraction limit. Science **1992**, *257* (5067), 189–195.
20. Dunn, R.C. Near-field scanning optical microscopy. Chem. Rev. **1999**, *99* (10), 2891–2927.
21. Mirkin, C.A.; Rogers, J.A. Emerging methods for micro-nd nanofabrication. MRS Bull. **2001**, *26* (7), 506–508.
22. Vettiger, P.; Despont, M.; Drechsler, U.; Durig, U.; Haberle, W.; Lutwyche, M.I.; Rothuizen, H.E.; Stutz, R.; Widmer, R.; Binnig, G.K. The "millipede"—More than one thousand tips for future AFM data storage. IBM J. Res. Develop. **2000**, *44* (3), 323–340.
23. Ulman, A. Self-assembled monolayers of alkyltrichlorosilanes—Building-blocks for future organic materials. Adv. Mater. **1990**, *2* (12), 573–582.
24. Bain, C.D.; Whitesides, G.M. Molecular-level control over surface order in self-assembled monolayer films of thiols on gold. Science **1988**, *240* (4848), 62–63.
25. Strong, L.; Whitesides, G.M. Structures of self-assembled monolayer films of organosulfur compounds adsorbed on gold single-crystals—Electron-diffraction studies. Langmuir **1988**, *4* (3), 546–558.
26. Zhao, B.; Brittain, W.J. Polymer brushes: Surface-immobilized macromolecules. Prog. Polym. Sci. **2000**, *25* (5), 677–710.
27. Prucker, O.; Rühe, J. Synthesis of poly(styrene) monolayers attached to high surface area silica gels through self-assembled monolayers of azo initiators. Macromolecules **1998**, *31* (3), 592–601.
28. Prucker, O.; Rühe, J. Mechanism of radical chain polymerizations initiated by azo compounds covalently bound to the surface of spherical particles. Macromolecules **1998**, *31* (3), 602–613.
29. Jordan, R.; Ulman, A.; Kang, J.F.; Rafailovich, M.H.; Sokolov, J. Surface-initiated anionic polymerization of styrene by means of self-assembled monolayers. J. Am. Chem. Soc. **1999**, *121* (5), 1016–1022.
30. Ingall, M.D.K.; Honeyman, C.H.; Mercure, J.V.; Bianconi, P.A.; Kunz, R.R. Surface functionalization and imaging using monolayers and surface-grafted polymer layers. J. Am. Chem. Soc. **1999**, *121* (15), 3607–3613.
31. Chen, W.; Fadeev, A.Y.; Hsieh, M.C.; Oner, D.; Youngblood, J.; McCarthy, T.J. Ultrahydrophobic and ultralyophobic surfaces: Some comments and examples. Langmuir **1999**, *15* (10), 3395–3399.
32. Husemann, M.; Mecerreyes, D.; Hawker, C.J.; Hedrick, J.L.; Shah, R.; Abbott, N.L. Surface-initiated polymerization for amplification of self-assembled monolayers patterned by microcontact printing. Angew. Chem. Int. Ed. **1999**, *38* (5), 647–649.
33. Prucker, O.; Naumann, C.A.; Rühe, J.; Knoll, W.; Frank, C.W. Photochemical attachment of polymer films to solid surfaces via monolayers of benzophenone derivatives. J. Am. Chem. Soc. **1999**, *121* (38), 8766–8770.
34. Gurtner, C.; Wun, A.W.; Sailor, M. Surface modification of porous silicon by electrochemical reduction of organo halides. Angew. Chem. Int. Ed. **1999**, *38* (13–14), 1966–1968.
35. Fürstner, A. Olefin metathesis and beyond. Angew. Chem. Int. Ed. **2000**, *39* (17), 3013–3043.
36. Grubbs, R.H.; Chang, S. Recent advances in olefin metathesis and its application in organic synthesis. Tetrahedron **1998**, *54* (18), 4413–4450.
37. Ivin, K.J.; Mol, J.C. *Olefin Metathesis and Metathesis Polymerization*; Academic Press: San Diego, 1997.
38. Matyjaszewski, K.; Xia, J.H. Atom transfer radical polymerization. Chem. Rev. **2001**, *101* (9), 2921–2990.
39. Patten, T.E.; Matyjaszewski, K. Copper(I)-catalyzed atom transfer radical polymerization. Acc. Chem. Res. **1999**, *32* (10), 895–903.
40. Liu, G.Y.; Amro, N.A. Positioning protein molecules on surfaces: A nanoengineering approach to supramolecular chemistry. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99* (8), 5165–5170.
41. Hyun, J.; Ahn, S.J.; Lee, W.K.; Chilkoti, A.; Zauscher, S. Molecular recognition-mediated fabrication of protein nanostructures by dip-pen lithography. Nano Lett. **2002**, *2* (11), 1203–1207.
42. Xu, S.; Liu, G.Y. Nanometer-scale fabrication by simultaneous nanoshaving and molecular self-assembly. Langmuir **1997**, *13* (2), 127–129.
43. Xu, S.; Miller, S.; Laibinis, P.E.; Liu, G.Y.

Fabrication of nanometer scale patterns within self-assembled monolayers by nanografting. Langmuir **1999**, *15* (21), 7244–7251.
44. Liu, G.Y.; Xu, S.; Qian, Y.L. Nanofabrication of self-assembled monolayers using scanning probe lithography. Acc. Chem. Res. **2000**, *33* (7), 457–466.
45. Sugimura, H.; Nakagiri, N. Scanning probe anodization: Nanolithography using thin films of anodically oxidizable materials as resists. J. Vac. Sci. Technol., A, Vac. Surf. Films **1996**, *14* (3), 1223–1227.
46. Legrand, B.; Stievenard, D. Nanooxidation of silicon with an atomic force microscope: A pulsed voltage technique. Appl. Phys. Lett. **1999**, *74* (26), 4049–4051.
47. Sugimura, H.; Takai, O.; Nakagiri, N. Multilayer resist films applicable to nanopatterning of insulating substrates based on current-injecting scanning probe lithography. J. Vac. Sci. Technol., B **1999**, *17* (4), 1605–1608.
48. Zamborini, F.P.; Crooks, R.M. Nanometer-scale patterning of metals by electrodeposition from an STM tip in air. J. Am. Chem. Soc. **1998**, *120* (37), 9700–9701.
49. Gorman, C.B.; Carroll, R.L.; He, Y.F.; Tian, F.; Fuierer, R. Chemically well-defined lithography using self-assembled monolayers and scanning tunneling microscopy in nonpolar organothiol solutions. Langmuir **2000**, *16* (15), 6312–6316.
50. Maoz, R.; Frydman, E.; Cohen, S.R.; Sagiv, J. "Constructive nanolithography": Inert monolayers as patternable templates for in-situ nanofabrication of metal–semiconductor–organic surface structures—A generic approach. Adv. Mater. **2000**, *12* (10), 725–731.
51. Maoz, R.; Frydman, E.; Cohen, S.R.; Sagiv, J. Constructive nanolithography: Site-defined silver self-assembly on nanoelectrochemically patterned monolayer templates. Adv. Mater. **2000**, *12* (6), 424–429.
52. Kolb, D.M.; Ullmann, R.; Will, T. Nanofabrication of small copper clusters on gold(111) electrodes by a scanning tunneling microscope. Science **1997**, *275* (5303), 1097–1099.
53. Hofmann, D.; Schindler, W.; Kirschner, J. Electrodeposition of nanoscale magnetic structures. Appl. Phys. Lett. **1998**, *73* (22), 3279–3281.
54. Randler, R.J.; Kolb, D.M.; Ocko, B.M.; Robinson, I.K. Electrochemical copper deposition on Au(100): A combined in situ STM and in situ surface x-ray diffraction study. Surf. Sci. **2000**, *447* (1–3), 187–200.
55. Schindler, W.; Hofmann, D.; Kirschner, J. Nanoscale electrodeposition: A new route to magnetic nanostructures? J. Appl. Phys. **2000**, *87* (9), 7007–7009.
56. Yang, R.; Evans, D.F.; Hendrickson, W.A. Writing and reading at nanoscale with a scanning tunneling microscope. Langmuir **1995**, *11* (1), 211–213.
57. Forouzan, F.; Bard, A.J. Evidence for faradaic processes in scanning probe microscopy on mica in humid air. J. Phys. Chem., B **1997**, *101* (50), 10876–10879.
58. Bard, A.J.; Denuault, G.; Lee, C.; Mandler, D.; Wipf, D.O. Scanning electrochemical microscopy—A new technique for the characterization and modification of surfaces. Acc. Chem. Res. **1990**, *23* (11), 357–363.
59. Cruchon-Dupeyrat, S.; Porthun, S.; Liu, G.Y. Nanofabrication using computer-assisted design and automated vector-scanning probe lithography. Appl. Surf. Sci. **2001**, *175*, 636–642.
60. Kenseth, J.R.; Harnisch, J.A.; Jones, V.W.; Porter, M.D. Investigation of approaches for the fabrication of protein patterns by scanning probe lithography. Langmuir **2001**, *17* (13), 4105–4112.
61. Piner, R.D.; Zhu, J.; Xu, F.; Hong, S.H.; Mirkin, C.A. "Dip-pen" nanolithography. Science **1999**, *283* (5402), 661–663.
62. Hong, S.H.; Zhu, J.; Mirkin, C.A. A new tool for studying the in situ growth processes for self-assembled monolayers under ambient conditions. Langmuir **1999**, *15* (23), 7897–7900.
63. Piner, R.D.; Mirkin, C.A. Effect of water on lateral force microscopy in air. Langmuir **1997**, *13* (26), 6864–6868.
64. Mirkin, C.A.; Hong, S.H.; Demers, L.M.; Zhu, J.; Eidelman, A.A. Multiple-ink-dip pen nanolithography. Abstr. Pap. Am. Chem. Soc. **2000**, *219*, U285.
65. Ahn, S.-J.; Lee, W.-K.; Zauscher, S. Fabrication of Stimulus—Responsive Polymeric Nanostructures by Proximal Probes. In *Bioinspired Nanoscale Hybrid Systems*; Schmid, G., Simon, U., Stranick, S.J., Arrivo, S.M., Hong, S., Eds.; Materials Research Society: Warrendale, PA, 2003; Vol. 735, C11.54.11–C11.54.16.
66. Kaholek, M.; Lee, W.-K.; LaMattina, B.; Caster, K.C.; Zauscher, S. Polymerization, Nanopatterning, and Characterization of Surface Confined, Stimulus-Responsive Polymer Brushes. In *Polymer Brushes: From Synthesis to Functional Microstructures*; Brittain, W.J., Advincula, R.C., Rühe, J., Caster, K.C., Eds.; John Wiley & Sons: New York, 2003. Submitted.
67. Demers, L.M.; Ginger, D.S.; Park, S.J.; Li, Z.; Chung, S.W.; Mirkin, C.A. Direct patterning of modified oligonucleotides on metals and insulators

by dip-pen nanolithography. Science **2002**, *296* (5574), 1836–1838.
68. Lee, K.B.; Park, S.J.; Mirkin, C.A.; Smith, J.C.; Mrksich, M. Protein nanoarrays generated by dip-pen nanolithography. Science **2002**, *295* (5560), 1702–1705.
69. Prime, K.L.; Whitesides, G.M. Self-assembled organic monolayers—Model systems for studying adsorption of proteins at surfaces. Science **1991**, *252* (5009), 1164–1167.
70. Hyun, J.; Lee, W.-K.; Nath, N.; Chilkoti, A.; Zauscher, S. Stimulus-responsive elastin-like polypeptide nanostructures. Nano Lett. **2003**. Submitted.
71. Lee, K.B.; Lim, J.H.; Mirkin, C.A. Protein nanostructures formed via direct-write dip-pen nanolithography. J. Am. Chem. Soc. **2003**, *125* (19), 5588–5589.
72. MacBeath, G.; Schreiber, S.L. Printing proteins as microarrays for high-throughput function determination. Science **2000**, *289* (5485), 1760–1763.
73. Li, Y.; Maynor, B.W.; Liu, J. Electrochemical AFM "dip-pen" nanolithography. J. Am. Chem. Soc. **2001**, *123* (9), 2105–2106.
74. Maynor, B.W.; Filocamo, S.F.; Grinstaff, M.W.; Liu, J. Direct-writing of polymer nanostructures: Poly(thiophene) nanowires on semiconducting and insulating surfaces. J. Am. Chem. Soc. **2002**, *124* (4), 522–523.
75. Wegscheider, S.; Kirsch, A.; Mlynek, J.; Krausch, G. Scanning near-field optical lithography. Thin Solid Films **1995**, *264* (2), 264–267.
76. Alkaisi, M.M.; Blaikie, R.J.; McNab, S.J. Nanolithography in the evanescent near field. Adv. Mater. **2001**, *13* (12–13), 877–887.
77. Inouye, Y.; Kawata, S. Near-field scanning optical microscope with a metallic probe tip. Opt. Lett. **1994**, *19* (3), 159–161.
78. Jersch, J.; Demming, F.; Hildenhagen, L.J.; Dickmann, K. Field enhancement of optical radiation in the nearfield of scanning probe microscope tips. Appl. Phys., A: Mater. Sci. Process. **1998**, *66* (1), 29–34.
79. Wurtz, G.; Bachelot, R.; H'Dhili, F.; Royer, P.; Triger, C.; Ecoffet, C.; Lougnot, D.J. Photopolymerization induced by optical field enhancement in the vicinity of a conducting tip under laser illumination. Jpn. J. Appl. Phys., Part 2-Lett. **2000**, *39* (2A), L98–L100.
80. Ecoffet, C.; Espanet, A.; Lougnot, D.J. Photopolymerization by evanescent waves: A new method to obtain nanoparts. Adv. Mater. **1998**, *10* (5), 411–414.

Polymer-Mediated Self-Assembly of Nanoparticles

Tyler B. Norsten
Amitav Sanyal
Roy Shenhar
Vincent M. Rotello
University of Massachusetts, Amherst, Massachusetts, U.S.A.

INTRODUCTION

Polymer self-assembly of nanoparticles yields stabilized metal and semiconductor nanocomposites as well as provides access to structured nanoscale materials. The properties of these materials can be tuned by varying the size and shape of the metal cluster as well as by altering the composition of the monolayer that covers the particles. Furthermore, the functionality and structure of the polymer can be modified, which lends an additional element of control to the entire assembly process. This "bottom-up" approach provides a methodology that is complementary to "top-down" lithographic methods, providing access to structures smaller and with greater 3-dimensional (3-D) control than is possible through sophisticated lithographic techniques such as electron-beam lithography.[1–4] Presently, one of the major obstacles facing nanotechnology is the successful integration of these two general approaches. Multiscale engineering of this sort would allow for the fabrication of intricate functional devices with atomic level structural control that manifests and spans itself into the macroscopic world.

One of the key initial goals in polymer-mediated nanoparticle assembly is the design of catalytic nanocomposites. We have therefore designated a section within this chapter solely to polymer–nanoparticle assemblies as applied to catalytic systems. The remaining sections describe the use of polymers to assemble nanoparticles into interesting and potentially useful architectures. Because of lack of space, this chapter is not an exhaustive review of the literature but instead contains selected examples, which we feel highlight many of the different methodologies that have been employed for the assembly of polymer-based nanoparticle composites. When necessary the reader will be directed to more in-depth articles and reviews that contain information pertinent to the area.

GENERAL APPROACHES TO POLYMER-MEDIATED ASSEMBLY OF NANOPARTICLES

Metal and semiconductor nanoparticles embedded within polymer matrices are attractive materials as they share the processing and handling advantages of bulk materials as well as unique possessing properties arising from the nanoscopic size of the components. Traditional preparation of polymer–nanoparticle composites has been accomplished by in situ generation of metal nanoclusters, typically by chemical reduction of a metal salt contained within a polymer matrix (Fig. 1, route A). The primary function of the polymer matrix is to provide a protective barrier against aggregation of the metal particles during this reduction process. As shown in Fig. 1, a variety of different polymers (homopolymers, block copolymers, and dendrimers) have been employed in this regard in an effort to create well-ordered nanocomposite materials for various applications.

A similar but separate approach involves the assembly of polymer scaffolds with preformed monolayer-protected clusters (MPCs). Monolayer-protected clusters already contain a surface monolayer that prevents the metal from agglomerating and allows for ease of handling and processing (Fig. 1, route B). As such, MPCs can be readily dispersed into polymer matrices taking advantage of designed interfacial interactions between the monolayer functional groups and the polymer backbone. Multiple groups have demonstrated over the past several years that the latter approach provides a versatile alternative method for the fabrication of novel nanoparticle-polymeric materials.

POLYMER SCAFFOLDS AND NANOPARTICLE BUILDING BLOCKS

Polymer Scaffolds

Polymers are large molecular structures in which many small repeat units are linked together. These repeat units can be linked together in a linear array, as is the case with the homopolymers and block copolymers described in Fig. 1, or in branched fashion as is the case with dendrimers. Many different routes were followed to synthetically prepare the various types of polymers described in this chapter; however, because of space limitations this is beyond the scope of the present chapter. Suffice it to say there has been significant recent progress

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013685

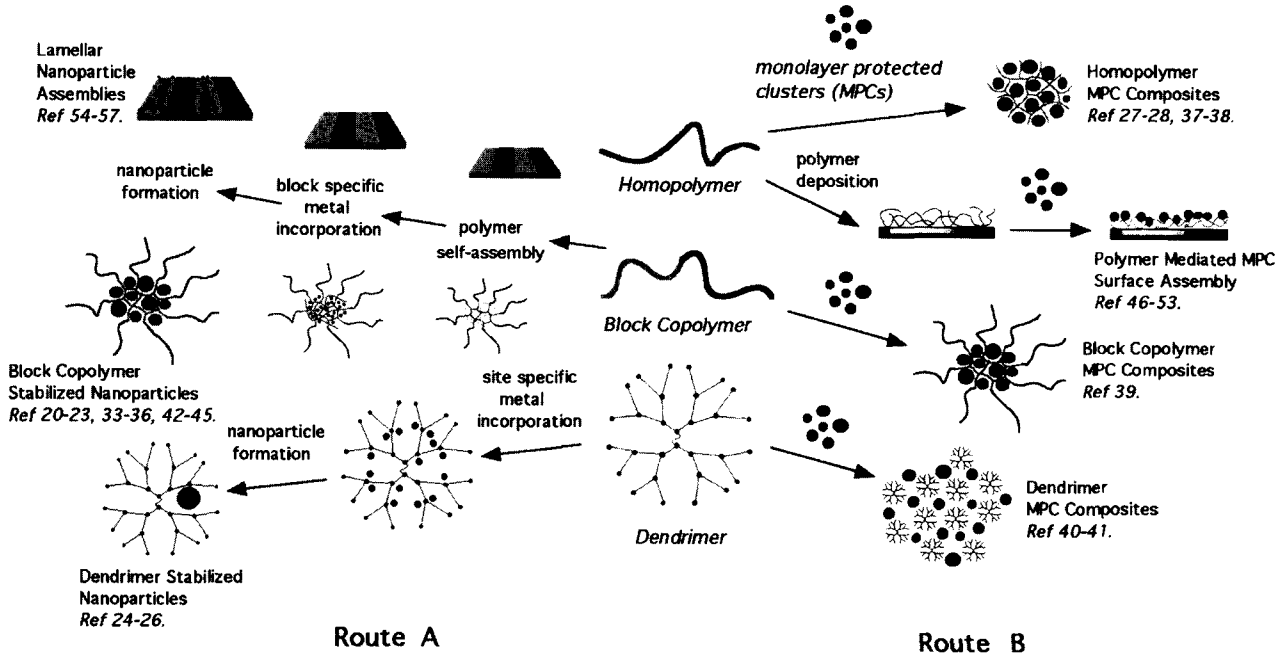


Fig. 1 Polymer-mediated approaches to fabrication of ordered nanocomposites. Route A: In situ formation of particles within polymer matrices. Route B: Polymer-mediated assembly of preformed monolayer protected clusters (MPCs).

in the area of polymer synthesis specifically with respect to the formation of block copolymers via living polymerization techniques.[5–7]

Block copolymers represent a new and significant class of materials because of their remarkable ability to form unique micro- and nanophase morphologies. These properties are unattainable with homopolymers or their corresponding polymer blends. Diblock copolymers differ from traditional homopolymers in that two chemically distinct polymer chains are covalently connected end-to-end. The immiscibility of the different blocks with each other or with a solvent causes these polymers to phase separate, as such diblock copolymers are able to self-assemble into solvent-stabilized micelles, vesicles, and regular surface stabilized lamellar, cylindrical, or spherical arrays. Several examples will illustrate how appropriately functionalized diblock copolymers are able to incorporate nanoparticles or their precursor metal salts into specific polymer blocks.

Dendrimers are hyperbranched polymers that differ from their traditional linear counterparts in that the repeating units cascade outwardly from a central core rather than elongating in a linear fashion.[8] Dendrimers with sufficient layers (generations) take on 3-D/globular morphologies containing interior voids and cavities capable of accommodating guest species. Functionality can be accurately placed throughout the core and at the periphery of dendrimers providing control over such features as site-specific molecular recognition events, interior cavity size, and overall polymer solubility. The internal attributes described above have been used to incorporate nanoparticles inside of dendrimers, while their distinct generational size-dependent morphologies have been used to define interparticle spacing between metal nanoparticles.

The use of biological polymers such as DNA to assemble and organize nanoparticles is an attractive strategy as the degree of noncovalent complementarity between these polymers can be effectively tuned. As the scope of nanoparticle–biopolymer interactions is exceedingly large in essence commanding its own chapter, we will not cover the topic here, there are, however, several recent examples worth noting that exemplify this type of assembly strategy.[9–11]

Nanoparticle Building Blocks

Fig. 1 describes two of the most typically followed routes to the formation of polymer–nanoparticle composites. Route A generally involves using the polymer as both a template for the formation of nanoparticles into nano- or microdomains defined by the polymer as well as an organic stabilizer to prevent agglomeration after nanoparticle formation. Fairly monodisperse nanoparticles can be assembled using this methodology; however, in many instances the supporting polymer matrix does not

offer long-term stability against agglomeration of the metal cores.

Route B typically involves the assembly of polymers with preformed nanoparticles already possessing a stabilizing monolayer. One of the most frequently encountered MPCs is that of the alkanethiolate-stabilized gold cluster synthesized by the solution-phase method developed by Brust et al.[12] (Fig. 2). In this procedure chemical reduction of a gold salt by a hydride-reducing agent in the presence of thiol capping ligands furnishes the desired monolayer-protected nanoparticles. By carefully controlling the reaction conditions and by varying the stoichiometry of the metal salt to capping ligand a variety of different core sizes (1.5–8 nm) can be obtained.

Surface modifications of these nanoparticles can be achieved through a place exchange reaction, where new incoming thiol ligands displace monolayer thiols on the parent MPC to afford mixed monolayer-protected clusters (MMPCs).[13] This monolayer modification technique is extremely versatile as it allows for the incorporation of virtually any functionalized thiol into the preexisting monolayer. This permits the interfacial interactions on the surface of the nanoparticle to be precisely tuned to that of the assembling polymer strands.

Gold MPCs are attractive building blocks for the fabrication of nanoscale architectures as their synthesis permits good control over their size and polydispersity. Furthermore, the alkanethiolate monolayer provides excellent stability to intercore agglomeration both in air and in solution. This permits the particles to be handled (i.e., dissolved in solvents, isolated, dispersed in matrices) multiple times without significant degradation occurring and allows for easy characterization using standard solution-phase techniques such as nuclear magnetic resonance (NMR), infrared (IR), and UV–VIS spectroscopy.[14]

There are many other types of metallic, bimetallic, and semiconductor nanoparticles available;[15,16] however, in most instances, the chemistries of their formation are quite harsh, which substantially reduces the variety of functionalities that can be placed in their protecting monolayers. This makes assembling these types of nanoparticles into extended ordered structures substantially more difficult. Nanoparticle chemistry is still very much in its infancy and it is only a matter of time before new procedures are created or existing procedures are modified to permit a wider variety of functionalities onto these other types of nanoparticles.

Monolayer-protected clusters clearly provide excellent "building blocks" of regular shape and size for the fabrication of larger ordered structures. These attributes provide part of the solution to the problem of multiscale fabrication as discussed earlier. What remains is their successful incorporation into desired functional ensembles.

ASSEMBLIES OF NANOPARTICLES FOR CATALYIC APPLICATIONS

Creation of nanoparticle assemblies possessing an open, porous framework is an attractive strategy for catalyst design. Various noble metals such as palladium, platinum, etc., are used for catalysis in numerous important chemical transformations, including hydrogenation, oxidation, and Heck coupling.[17–19] A high surface area-to-volume ratio is desirable for the creation of efficient heterogeneous catalysts as the cost of these precious metals is significant. Small nanoparticles (1–4 nm) are excellent candidates for catalysts in this respect as they exhibit extremely high surface area-to-volume ratios by virtue. Catalytically active materials obtained by metal nanoparticles dispersed in polymeric matrices display increased stability, improved processability, recyclability, and solubility in a variety of organic solvents.

A major thrust in the area of nanoparticle-based catalysts has been the preparation of palladium colloids within the micelles of amphiphilic block copolymers.[20] Forster and Antonietti have utilized poly-4-vinylpyridine-*b*-polystyrene (P4VP-*b*-PS) as the block copolymer and

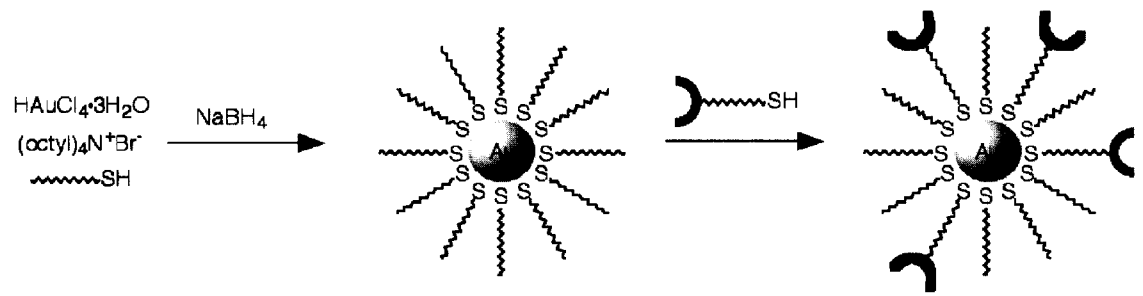


Fig. 2 Brust et al.'s solution phase MPC synthesis,[12] followed by MMPC synthesis using the place exchange process developed by Murray.[13]

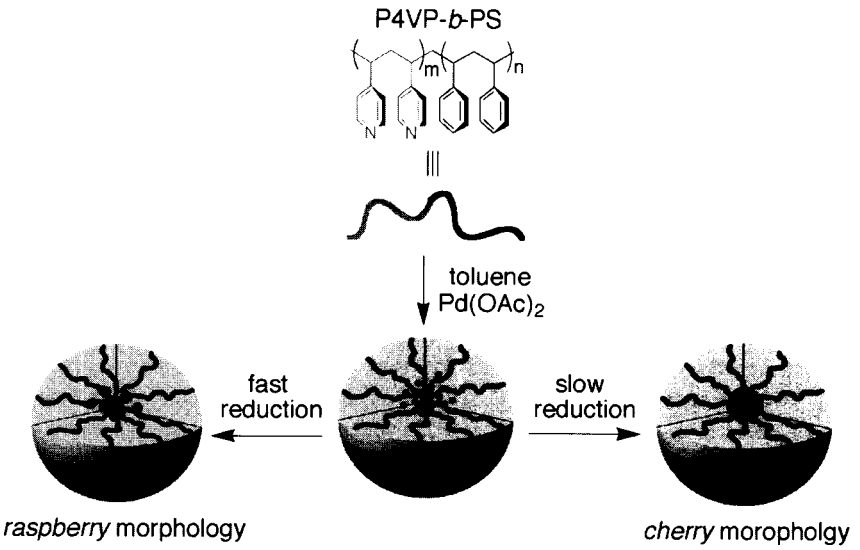


Fig. 3 Preparation of palladium colloids in P4VP-*b*-PS block copolymer micelles.

$Pd(OAc)_2$ as the palladium source in the preparation of metal colloids (Fig. 3). Distinct colloidal morphologies can be obtained by carefully controlling the method of reduction: A "raspberry morphology" is created with rapid reduction vs. a "cherry morphology" with a slower reduction rate.[21] Colloidal aggregates with the "raspberry morphology," where the particle size is smaller, demonstrated excellent catalytic activity in the carbon–carbon bond forming Heck reaction. Although the catalyst reactivities were similar to other widely used phosphine ligand/Pd complexes, the polymer-bound Pd nanoparticle catalyst showed much higher stability.

Li and coworkers have shown that palladium nanoparticles stabilized by poly(*N*-vinyl-2-pyrrolidone) (PVP) can be used as efficient catalysts for Suzuki coupling reactions in 40% ethanolic aqueous media.[22] Good conversions were obtained upon refluxing for 12–48 hr; however, the PVP homopolymer does not provide adequate stability under such harsh reaction conditions as precipitation of the palladium metal occurs at the end of the reaction.

Underhill and Liu have employed triblock copolymer cross-linked nanospheres as templates for Pd nanoparticles.[23] Micelles of a polyisoprene-*b*-poly(2-cinnamoyloxyethyl methacrylate)-*b*-poly(*tert*-butyl acrylate) (PI-*b*-PCEMA-*b*-P*t*BA) were formed in 65% hexane/THF mixture. The micelles consisted of PI coronas, PCEMA shells, and P*t*BA cores. Photo-cross-linking the PCEMA shell produced stable nanospheres, hydroxylation of the terminal isoprene units provided water-dispersibility, and subsequent hydrolysis of the P*t*BA groups provided poly(acrylic acid) (PAA) cores capable of complexing $PdCl_2$. Loading of $PdCl_2$ into the acrylic acid core followed by reduction using hydrazine yielded polymer-stabilized palladium nanoparticles (Fig. 4). These palladium-loaded nanospheres were used to study hydrogenation reactions of tetraethylammonium bromide, vinylacetic acid, methylmethacrylate, and ethylene glycol dimethylacrylate. Although the diffusion barriers created by the encapsulating nanospheres led to slightly lower reaction rates compared to palladium black, the hydrogenation kinetics can be controlled by changing the pH as the core structure of the nanospheres is pH sensitive, which ultimately affects the transport of charged substrates into the nanospheres.

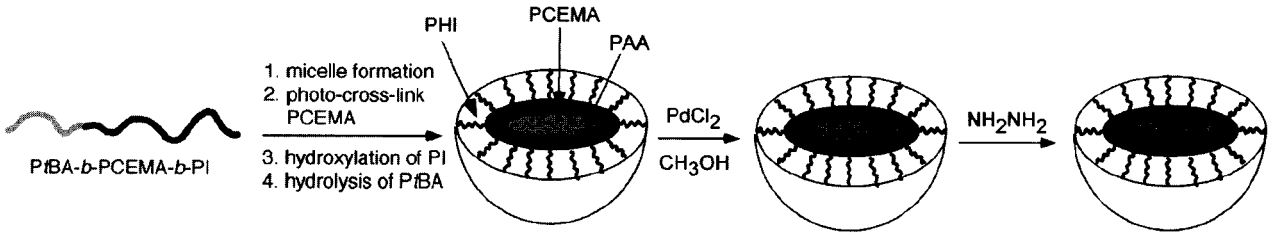


Fig. 4 Cross-linked triblock copolymer nanospheres encapsulating Pd nanoparticles.

Dendrimers have also been used to sequester nanoparticles within their hyperbranched polymer architectures. Crooks et al. have employed poly(amidoamine) (PAMAM) and poly(propylene imine) (PPI) dendrimers as templates for the in situ generation of nanoparticles within the polymer shell of the dendrimer in an effort to create catalytically active nanocomposites.[24] Monodisperse nanoparticles can be created inside the dendrimer shell by first introducing a metal salt that strongly complexes to the interior tertiary amine groups followed by subsequent reduction of the metal ions producing a trapped-stabilized zero-valent metal nanoparticle. They were able to demonstrate size selective hydrogenation catalysis of α-substituted allyl alcohols by employing Pd stabilized by hydroxy-terminated PAMAM dendrimers (Fig. 5).[25] As generations were added to the dendrimers (G4OH→ G6OH→G8OH) turnover frequencies of all the substrates decreased. Furthermore, within a given generation of dendrimer the turnover frequencies were always smaller for the larger substrates. The observed selectivity is a result of the steric crowding at the periphery of the dendrimer, which increases with the dendrimers generation. In essence, the periphery of the dendrimer behaves like a molecular filter which can be tuned in a generational fashion, ultimately hampering access of larger substrates to the catalytically active site.

Recently, Mecking and coworkers have utilized amphiphilic hyperbranched poly(glycerols) to prepare catalytically active and stable palladium colloids.[26] This amphiphilic *dendrimer-like* molecule is capable of solubilizing precursor palladium salts (e.g., $PdCl_2$, $Pd(OAc)_2$) in nonpolar solvents such as toluene and chloroform. Slow reduction using hydrogen results in stable colloidal palladium. The authors suggest that the *soft* binding polyglycerols tend to avoid any ligand poisoning that may result in the amine-functionalized systems because of stronger coordination between the metal and the polymer–ligand. Subsequent catalytic hydrogenation reactions on cyclohexene using this amphiphile-stabilized palladium colloid proceeded at an impressive rate of 700 turnovers hr^{-1}, with no significant decrease in rate during recyclable use.

In the example shown above, Crooks et al. employed the hindering effect of a dendrimer sheath in an advantageous way to obtain size selective catalytic discrimination. The polymer casing effectively controls substrate access to the catalytically active site. In general, however, this tends to be a drawback of polymer-supported catalysts as the substrates must find their way through the stabilizing polymer shell in order to reach the buried metal catalyst. This ultimately results in inefficient mass transport mechanisms for these types of catalysts.

In an alternative approach to the polymer-stabilized nanoparticle catalysts described above, we employ directed self-assembly strategies to develop porous catalyst systems. Mixed monolayer-protected clusters of catalytically active metals can be utilized as building blocks in the presence of a suitable polymeric mediator to obtain macromolecular assemblies. Subsequent removal of the stabilizing polymers and ligands through calcination provides materials with tailored catalytic properties and with improved access to the catalytically active sites.

Preliminary studies to investigate this strategy involved a polymer-mediated three-component system employing both carboxylic acid-terminated gold and SiO_2 nanoparticles, and an amine-functionalized polystyrene random copolymer.[27] This type of assembly strategy is based on acid–base chemistry resulting in the electrostatic attraction between the basic polymer and the acidic nanoparticles when combined. Well-integrated nanocomposites were obtained upon the addition of the polymer to a mixture of the two nanoparticles, whereas segregated clusters were formed upon premixing of one of the MPCs with the polymer followed by the addition of the other nanoparticle.

This strategy was then applied to fabricate high-efficiency palladium catalysts.[28] Preformed SiO_2–COOH/poly-NH_2 was used as a scaffold to assemble

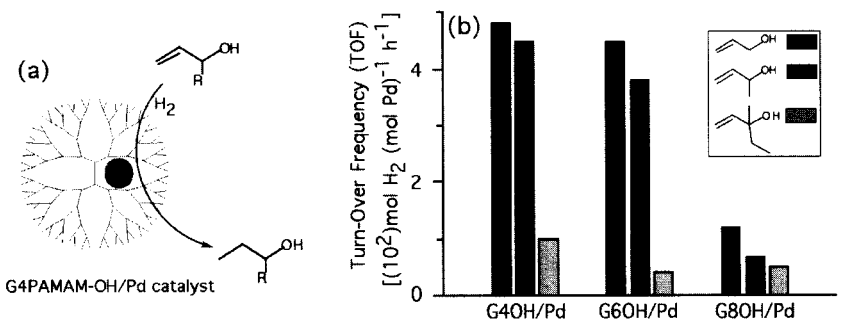


Fig. 5 (a) Representation of a PAMAM-stabilized Pd nanoparticle catalyzing the hydrogenation of α-substituted allyl alcohols. (b) Turnover frequencies of the hydrogenation reactions of various allyl alcohols with generation 4-, 6-, and 8-hydroxy-terminated dendrimers.

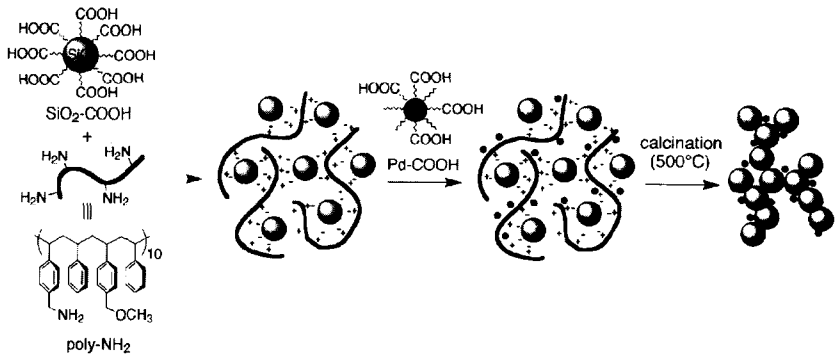


Fig. 6 Formation of the SiO_2-stabilized Pd nanoparticle catalyst via the three component self-assembly strategy.

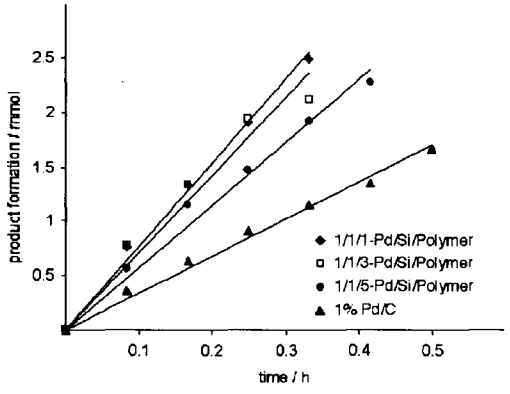
catalytically active Pd–COOH nanoparticles. The two-phase assembly strategy was necessary to obtain a high exposure of the catalytic metal on the surface of the final aggregate. Calcination removed all of the organic matter creating a highly porous Pd–SiO_2 composite material free of any polymer and monolayer stabilizers (Fig. 6).

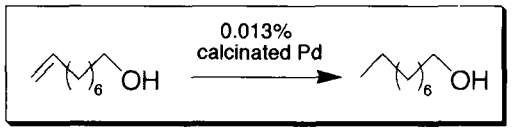
The catalytic activity of the calcinated aggregates was investigated for the hydrogenation of 9-decen-1-ol. High turnover frequencies (TOF) (10,100 hr^{-1} for the 1:1:1 w/w/w composition of Pd/Si/polymer, respectively) were observed, compared to the 7200 hr^{-1} found under the same conditions for the commercial 1% Pd/C catalyst (Fig. 7). In addition, these systems proved to be excellent catalysts for Heck coupling reactions between electronically activated bromoarenes and styrene or methyl acrylate at very low catalyst loadings (0.045 mol% of Pd).

Unlike commonly used homogeneous Pd-catalysts, all of the nanoparticle-based catalysts described above required no activation by additional ligands such as phosphines and most could be recycled with only small decreases in their activity. Furthermore, nanoparticle catalysts are much simpler both to separate from their reaction mixtures and to recycle than typical homogeneous catalysts. The combination of these attributes makes them particularly attractive and useful catalysts for a variety of chemical transformations.

ORGANIZED NANOPARTICLE ASSEMBLIES

The creation of discrete organized nanoparticle assemblies provides a structural motif for nanotechnology.[29–31] Highly structured, 3-D nanocomposites display unique magnetic, electronic, and optical behaviors that are both size and distance dependent. An effective way to control both the size and assembly of metallic nanoparticles is through the use of diblock copolymers. As discussed earlier, diblock copolymers consist of two chemically distinct segments capable of organizing into various structural morphologies because of microphase separation. The size and morphology of the domains are easily controlled by varying the lengths of the individual blocks.[32] These





Catalyst (Pd:Si:Polymer)	TOF/h^{-1}[a]
1:1:1	10154
1:1:3	9419
1:1:5	7622
1% Pd/C	7202

[a] calculated as $mol_{product} mol_{Pd}^{-1} h^{-1}$

Fig. 7 Rates of product formation in hydrogenation reactions of 9-decen-1-ol with different polymer ratios.

self-assembled structures can act as scaffolds for patterning with metallic nanoparticles, employing selective wetting or complexation of the polymeric domains with metal clusters.

Rutnakornpituk and coworkers were able to prepare superparamagnetic cobalt nanoparticles within poly[dimethylsiloxane]-*b*-poly[(3-cyanopropyl)methylsiloxane-]-*b*-poly[dimethylsiloxane] (PDMS-PCPMS-PDMS) triblock copolymer templates.[33] The copolymer forms micelles in toluene with the PCPMS block being oriented toward the interior of the micelle. Introduction of the cobalt precursor ($Co_2(CO)_8$) into the PCPMS block followed by subsequent thermal decomposition of the metal carbonyl complex yields stable nanoparticles. It is thought that the nitrile groups adhere to the particle surface while the bulky PDMS endblocks provide the steric stability against the core-to-core agglomeration.

Moore and coworkers synthesized several metal nanoparticle clusters using ionomers as templates.[34] The size of such ionic microdomains in these ion-containing copolymers can be controlled by choice of structure of the ionic monomer, whereas in diblock ionomers with spherical microdomains, the radii of the ionic core can be varied between 15 and 100 nm by varying the ionic block length. Utilizing these as microcompartments for nanocluster synthesis allows control over their size. In this approach the control of aggregate size is achieved primarily through ionomer aggregate size and not through varying ion concentration and subsequent annealing conditions.

Ng Cheong Chan et al. have demonstrated that block copolymers created from living ring opening metathesis polymerization reactions of methyltetracyclododecene (MTD) and metal-functionalized norbornene monomers can yield stable nanoparticles. They have shown that both a precursor metal complex (Ag(Hfacac)(COD))[35] as well as preformed monolayer-stabilized CdSe[36] nanoparticles can be incorporated into MTD and phosphine-functionalized norbornene block copolymers. In the case of the silver nanoparticles, the silver salts coordinate to the phosphine ligands within the core of solvent-stabilized micelles. Heating the silver-loaded micelles in water produces single nanoclusters (Fig. 8). Both fabrication methods yield stable spherical nanocomposites, but the latter method (akin to route B in Fig. 1) allows for the introduction of preformed nanoparticles that display a much narrower particle size distribution.

We have developed a different approach for using polymer scaffolds to direct the assembly of preformed nanoparticles. Our approach, which we have coined "bricks-and-mortar" self-assembly, involves functionalized nanoparticles as the building blocks or "bricks," and polymer scaffolds bearing complementary recognition units that serve as the "mortar." The complementary recognition motif is based on the three-point hydrogen bond between thymine-functionalized gold nanoparticles (Au-thy) and diaminotrazine-functionalized polystyrene (poly-triaz) (Fig. 9).

Addition of a triazine-substituted random copolymer (poly-triaz) to concentrated solutions of thymine-functionalized gold nanoparticles (Au-thy) in nonpolar solvents results in the formation of insoluble aggregates.[37] The TEM images of the THF-soluble fraction of the poly-triaz/Au-thy precipitate revealed the formation of large spherical clusters approximately 100 nm in diameter composed of 3000–7000 individual gold particles (Fig. 10). Small-angle X-ray scattering (SAXS) analysis of the aggregates suggests that the nanoparticles are regularly dispersed within the polymer matrix providing evidence that individual polymer strands weave the nanoparticles together. When hydrogen bonding is disrupted, as is the case with the methylated Au-thy nanoparticles, no such aggregates are formed. This provides

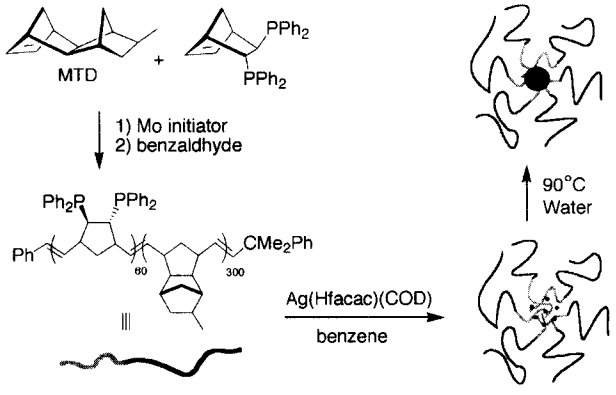


Fig. 8 Phosphine-containing diblock copolymer (bottom left) forming silver-loaded micelles (bottom right). Subsequent thermal treatment leads to stabilized individual silver nanoclusters (top right).

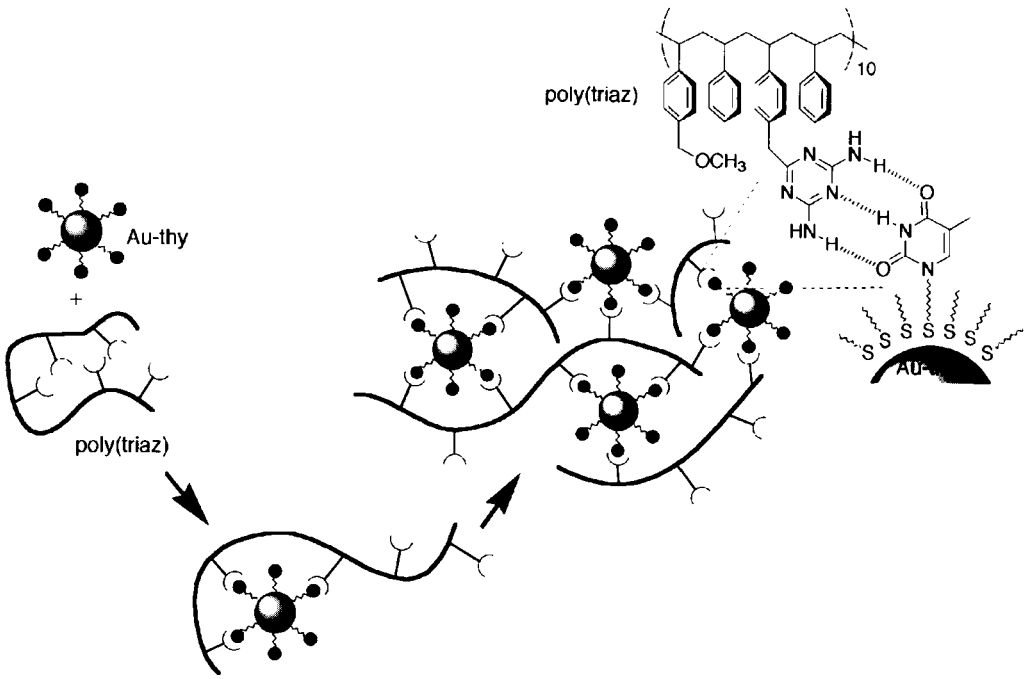


Fig. 9 Three-point hydrogen bonding recognition between thymine-capped gold nanoparticles and diaminotriazine-functionalized polymers leads to extended network formation.

direct evidence that the specific hydrogen bonding interactions between the polymer and the nanoparticle drive the assembly process.

Thermal control of the assembly process leads to aggregates of well-defined size and morphology. For instance, highly size-dispersed networks of spherical aggregates were obtained by performing the assembly at 10°C, while performing the aggregation at −20°C yields even larger, individual aggregates. These aggregates are 5–10 times larger than the aggregates formed at 23°C, ranging from 0.5 to 1 μm in diameter and composed of 0.6–5.0 million (!) individual Au-thy nanoparticles, and are among the most complex synthetic self-assembled structures ever constructed.

A follow-up study was undertaken aimed at understanding how secondary molecular interactions such π-π stacking in combination with the primary three-point hydrogen bonding interaction would influence the overall aggregation process.[38] Here various combinations of nanoparticle and polymer were synthesized containing different recognition elements consisting of either an electron-rich anthracene moiety or an electron-poor

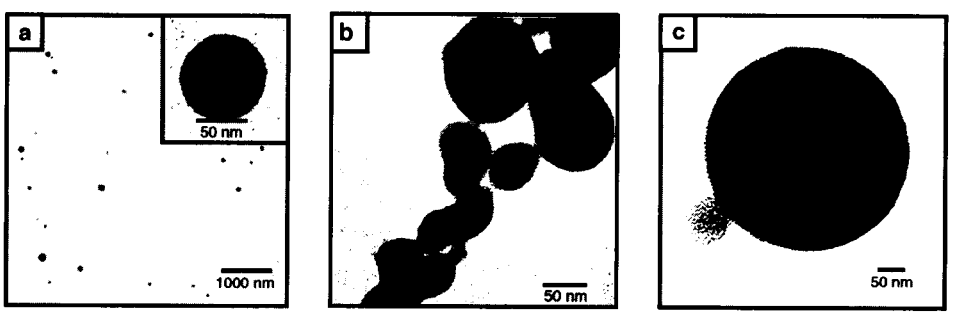


Fig. 10 (a) TEM of Au-thy/poly-triaz aggregates formed at 23°C. Inset: Representative self-assembled nanoparticle-polymer microsphere. (b) TEM of Au-thy/poly-triaz aggregate network formed at 10°C. (c) TEM of Au-thy/poly-triaz spherical aggregate formed at −20°C.

napthalimide moiety in conjunction with a primary hydrogen bonding element. This study showed that weaker secondary interactions are indeed important to aggregate formation and that overall aggregate morphology is directly related to supramolecular events that occur at the colloid–polymer interface.

While we were able to demonstrate a limited level of control over the size and the morphology using monoblock copolymers by carrying out the assembly at different temperatures, we were never able to fully rationalize the observed trends. A solution to this problem is to employ a diblock copolymer where one of the blocks is functionalized as before and the second block remains unfunctionalized.[39] In this approach the functionalized block will provide the required hydrogen bonding interactions for the assembly, while the inert second block will prevent extensive network formation. As only the functionalized block is capable of incorporating nanoparticle, the overall core size should depend primarily on the length of the functionalized block (Fig. 11a).

The feasibility of diblock-based control of aggregate size was demonstrated by employing three symmetric diblock copolymers (Fig. 11b) with varying total lengths. Combining Au-thy with the various polymers (a–c) results in aggregates with core sizes of 13.2±1.3, 13.9±1.2, and 19.4±1.8 nm, respectively, as revealed by TEM analysis. Dynamic light scattering (DLS) experiments of the polymer micelles reveal the effective hydrodynamic radii R_h (core and polystyrene corona) to be 18.7, 27.1, and 31.1 nm for the polymers (a–c), respectively. Comparison of core size by TEM to the overall aggregate size observed by DLS suggests that the polymer segments within the core are extended as compared to the polystyrene segments lying in the exterior, as the corona contributes less than half of the overall radius (Fig. 11c).

The ability to control the spacing between nanoparticles is an important issue as the electronic, and in certain cases the magnetic, properties of the nanoclusters are affected by neighboring particles. We were able to employ the size-dependent properties displayed by different generations of PAMAM dendrimers to both assemble and control the distance between gold nanoparticles.[40] Electrostatic self-assembly between the dendrimer and nanoparticle components was provided by salt-bridge formation between a carboxylic acid-functionalized nanoparticle and the terminal amine groups on PAMAM dendrimers (Fig. 12).

Examination of the resulting assemblies by TEM revealed that varying the ratios of MMPC to dendrimer resulted in two distinct morphologies. Low dendrimer to

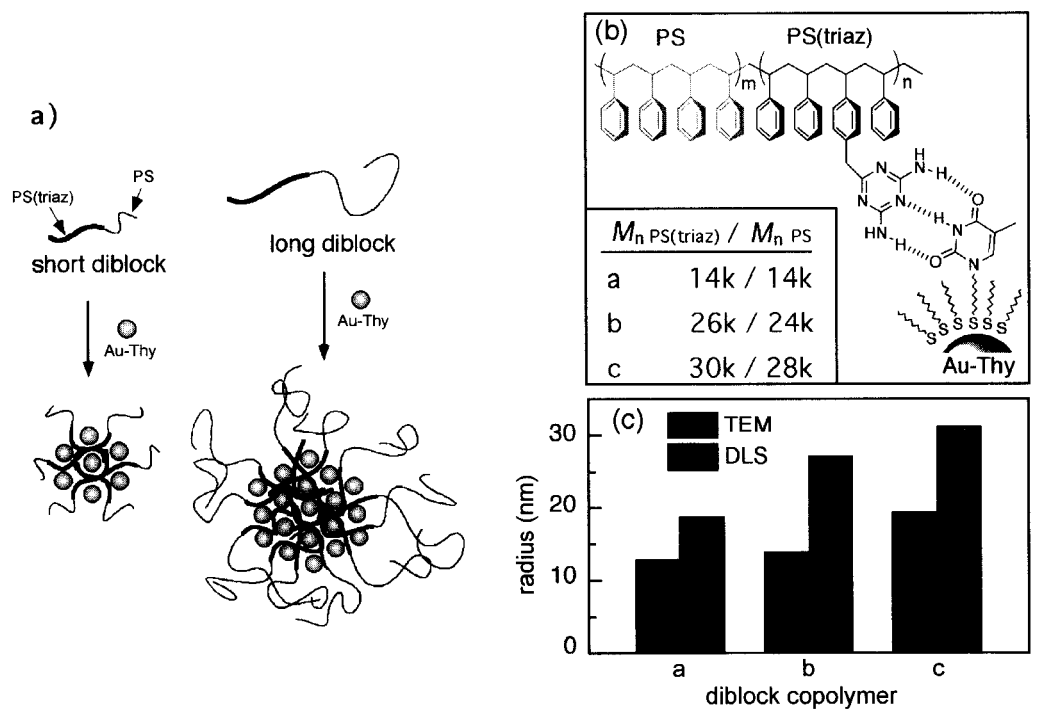


Fig. 11 (a) Schematic representation of diblock length controlling nanoparticle core and corona size. (b) Different length PS(triaz)-*b*-PS polymers assembled with Au-Thy. (c) Observed nanoparticle core size by TEM compared to hydrodynamic radii of the micelles determined by DLS.

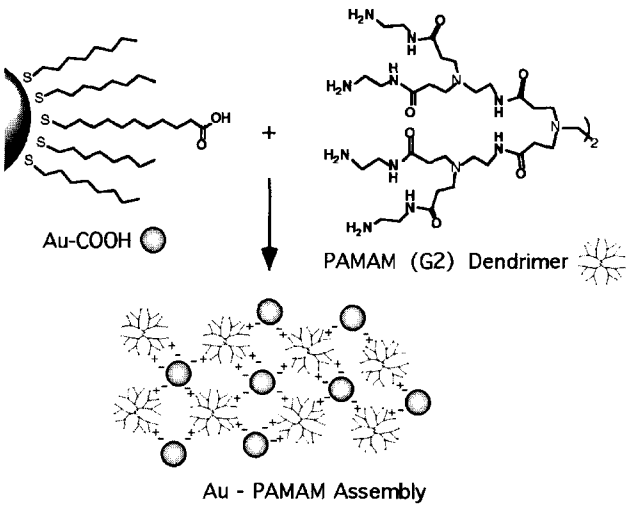


Fig. 12 Schematic representation of electrostatic self-assembly of carboxylic acid gold nanoparticles and PAMAM dendrimers.

MMPC ratios led to network-like structures, which showed little evidence of spatial control. In contrast, high ratios of dendrimer to MMPC (10:1 based on functional group equivalence) resulted in spherical aggregates ranging from 0.5 to 1.5 μm in diameter (Fig. 13a). In these assemblies, qualitative differences in inter-MMPC spacing could be readily discerned.

Quantification of interparticle distance in the dendrimer MMPC nanocomposites was obtained using SAXS. When the carboxylic acid-functionalized nanoparticle was assembled with PAMAM dendrimers, the primary peak shifted to lower q values (translating into greater interparticle distances) as the particle is assembled with larger dendrimers (G0→G1→G2→G4→G6), (Fig. 13). This example clearly demonstrates how different generations of dendrimers can be effectively used to control the spacing between nanoparticles.

Dendrimer–nanoparticle composites have also been used to form functional sensors.[41] Vossmeyer and coworkers created thin films of chemically cross-linked gold nanoparticles and polyphenylene dendrimers and deposited them onto glass substrates containing interdigitated gold electrodes. The rigid polyphenylene dendrimer provides a highly porous film for the incorporation of analyte while the nanoparticles provide a high surface-to-volume ratio leading to excellent signal transduction through the film. They were able to show ppm sensitivity to various volatile organic compounds (VOCs) such as toluene and tetrachloroethylene by measuring the change in electrical resistance through the films after exposure to the VOCs. This example demonstrates how the directed assembly of polymers and nanoparticles can lead to working composite devices.

HIGHLY ORGANIZED POLYMER–NANOPARTICLE ASSEMBLIES ON SURFACES

The previous examples show how polymers have been used to both create finite ordered nanoparticle arrays and assemble nanoparticles into discrete ordered entities extending into the microscale realm. The formation of these large complex structures from small polymolecular building blocks is a testament to the power of self-assembly. For many purposes, the controlled ordering at this level may be adequate; nevertheless, functional device applications such as nanoelectronics or memory media (optical or magnetic) require higher degrees of ordering. To create such sophisticated devices it is essential to fabricate not only highly *ordered* arrays but also arrays that are highly *extended* over vast distances (nm→cm).

To this end, Spatz and coworkers have used micelles obtained from diblock copolymers of polystyrene-polyvinylpyridine (PS-*b*-PVP)[42,43] and polystyrene-polyethylene oxide (PS-*b*-PEO)[44] for synthesis of metallic and semiconductor nanoparticles. In the case of the PS-*b*-PVP, the pyridine units within the miceller cores can uptake transition metal salts because of hard acid/soft base interactions (Fig. 14). This strategy allows one to obtain size-controlled synthesis of gold nanoparticles by controlling the amount of the precursor salt. Micelles containing the gold salt arranged into quasi-hexagonal packing when transferred onto a flat surface. Subsequent treatment with oxygen plasma completely removed all of the polymeric material yielding highly ordered bare gold nanoclusters.[45] Reducing the gold salts in the micellar core before plasma treatment gave clusters of similar size. Varying the concentration of metal salt allows one to obtain cluster sizes from 1 to 15 nm, whereas changing the length of the blocks allows one to get interparticle distances from 30 to 140 nm.

Using a different approach, similar to that of the layer-by-layer (LBL) assembly of oppositely charged polyionophores developed by Decher,[46] Schmid and coworkers

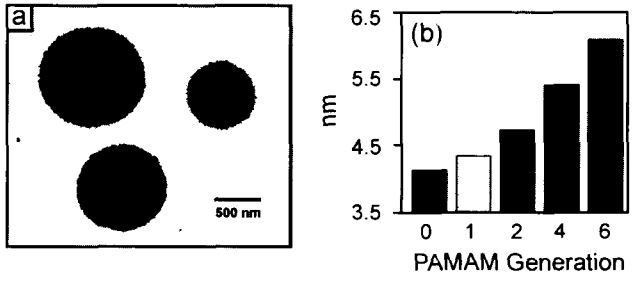


Fig. 13 (a) Self-assembled nanoparticle-dendrimer microspheres. (b) Graph of correlation between PAMAM generation and interparticle spacing derived from SAXS data.

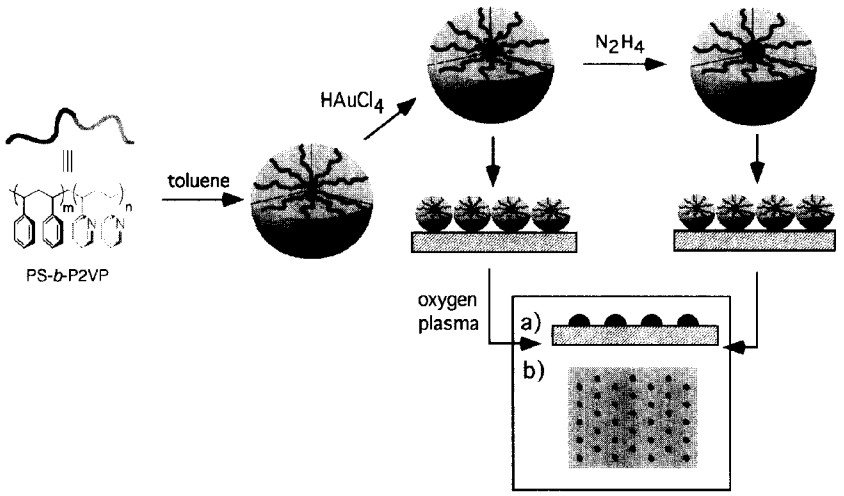


Fig. 14 Deposition of nanoparticles in a block copolymer micelle onto a surface. (a) Subsequent removal with oxygen plasma yields bare nanoparticles. (b) Schematic representation showing hexagonal packing arrangement of nanoparticles.

were able to assemble $[Au_{55}(Ph_2PC_6H_4SO_3H)_{12}Cl_6]$ clusters onto poly(ethyleneimine) (PEI)-covered surfaces (Fig. 15).[47] They found that the clusters ordered on the surface into the typical hexagonally close-packed arrangement, but they also found large domains of nanoparticles with an unusual cubic lattice structure. They attribute this high level of surface ordering to the use of PEI with a moderate molecular weight (60,000), which they believe allows the polymer to uncoil on the surface creating a more ordered polymer template. In this case the nanoparticles are packed immediately next to one another as there is no thick polymer sheath surrounding the nanoparticle. There are several examples of LBL assemblies containing multiple layers of polymers with gold,[48–51] semiconductor,[52] or magnetic nanoparticles,[53] and the interested reader is directed to those references.

Using the MTD and phosphine-functionalized norbornene block copolymers as described earlier (Fig. 8), Ng Cheong Chan et al. were able to form ordered surface assemblies of silver and gold nanoparticles.[54] In these studies the phosphine-containing block of the copolymer can selectively coordinate, in a ''dative'' fashion, to gold and silver coordination complexes in benzene solutions. Films cast from the block copolymer silver solutions yielded lamellar structures while those cast from the gold solutions gave cylinders. Thermally annealing the films yielded striped domains of silver, forming nanoparticles (2–10 nm in diameter) and spherical domains of gold nanoparticles (1.5–4 nm). In both cases long-range ordering occurred and the interdomain spacing was consistent with the length of the unfunctionalized block. This type of methodology was extended to the incorporation of zinc

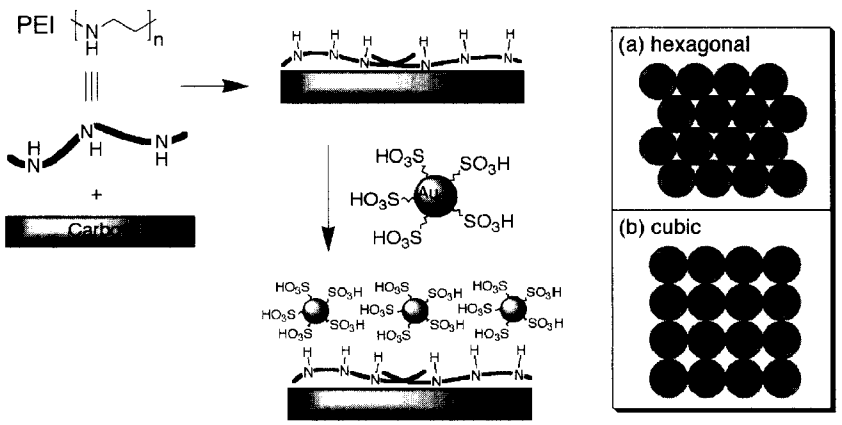


Fig. 15 Poly(ethyleneimine)-mediated assembly of sulfonated gold nanoparticles arranging into (a) hexagonal and (b) cubic-packed formations.

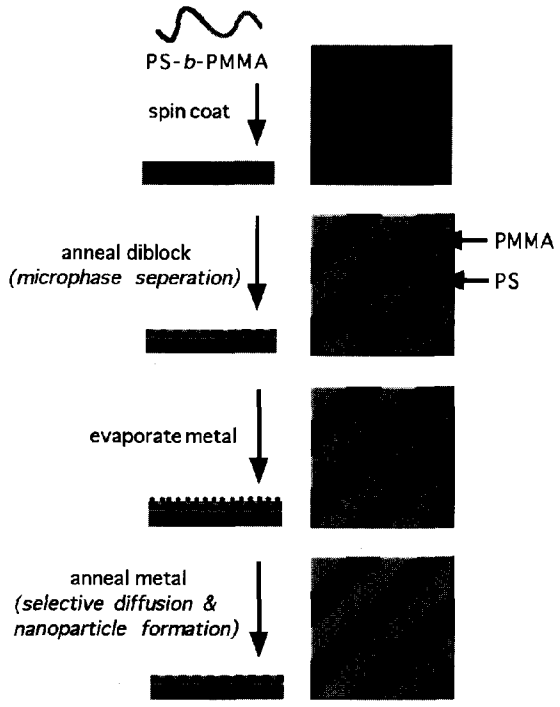


Fig. 16 Steps for creating ordered arrays of nanoclusters and nanowires in phase-separated PMMA-*b*-PS diblock copolymer films.

and cadmium into similar norbornene block copolymers to form the sulfides of these metals within specific lamellar block copolymer microdomains.[55]

More recently, Lopes and Jaeger have demonstrated that ultrathin films of phase-separated diblock copolymers can be utilized as templates allowing for the incorporation of metal nanoclusters with selectivity approaching 100% into the desired domain.[56,57] In these studies metal vapors were allowed to permeate an ultrathin film consisting of laterally alternating domains of PS and PMMA, obtained using a PS-*b*-PMMA diblock copolymer (Fig. 16). The PS domains preferentially absorb Au and Ag, while the PMMA domains show preference for In, Pb, Sn, and Bi. Near-perfect selective absorption is observed for the deposition of Ag. In the cases where the deposition is only partially selective to one block, selectivity is obtained by annealing the system above the glass transition temperature of the diblock copolymer. Depending on the annealing protocol, distinct separate nanoparticles could be formed within the decorated polymer block. Alternatively, metal nanowires can be obtained by a repeated deposition and annealing sequence. It was shown that the *nanowire* diblock templates display linear current–voltage (*I–V*) characteristics over a wide range of conditions, while the *nanoparticle* diblocks exhibit highly nonlinear *I–V* curves.

CONCLUSION

Polymer-mediated assembly of nanoparticles is a versatile and effective method for the creation of nanocomposite materials. The ability to chemically control both the polymer structure and the nanoparticle core and shell at the molecular level provides access to a wide variety of resulting materials with tunable properties. This combinatorial approach, coupled with the wealth of structures that can be obtained through self-assembly, makes polymer–nanoparticle composites potentially useful for a variety of applications. Nanoparticle chemistry and polymer self-assembly are still emerging disciplines and as such we are just beginning to learn the rules that govern their assembly. As we learn these rules and begin to put them to the test, both the complexity and the resulting utility of nanoparticle–polymer composites will increase dramatically.

REFERENCES

1. Jager, E.W.H.; Smela, E.; Inganas, O. Microfabricating conjugated polymer actuators. Science **2000**, *290*, 1540–1545.
2. Chen, Y.; Pepin, A. Nanofabrication: Conventional and nonconventional methods. Electrophoresis **2001**, *22*, 187–207.
3. Craighead, H.G. Nanoelectromechanical systems. Science **2000**, *290*, 1532–1535.
4. Wallraff, G.M.; Hinsberg, W.D. Lithographic imaging techniques for the formation of nanoscopic features. Chem. Rev. **1999**, *99*, 1801–1821.
5. Matyjaszewski, K.; Xia, J. Atom transfer radical polymerization. Chem. Rev. **2001**, *101*, 2921–2990.
6. Hawker, C.J.; Bosman, A.W.; Harth, E. New polymer synthesis by nitroxide mediated living radical polymerizations. Chem. Rev. **2001**, *101*, 3661–3688.
7. Hadjichristidis, N.; Pitsikalis, M.; Pispas, S.; Iatrou, H. Polymers with complex architecture by living anionic polymerization. Chem. Rev. **2001**, *101*, 3747–3792.
8. Newkome, G.R.; Moorefield, C.N.; Vögtle, F. *Dendritic Molecules—Concepts, Syntheses, Perspectives*; VCH: Weinheim, 1996.
9. Cao, Y.C.; Jin, R.; Mirkin, C.A. Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection. Science **2002**, *297*, 1536–1540.
10. Mucic, R.C.; Storhoff, J.J.; Mirkin, C.A.; Letsinger, R.L. DNA-directed synthesis of binary nanoparticle network materials. J. Am. Chem. Soc. **1998**, *120*, 12674–12675.

11. Mann, S.; Shenton, W.; Li, M.; Connolly, S.; Fitzmaurice, D. Biologically programmed nanoparticle assembly. Adv. Mater. **2000**, *12*, 147–510.
12. Brust, M.; Walker, M.; Bethell, D.; Schiffrin, D.J.; Whyman, R. Synthesis of thiol-derivatized gold nanoparticles in a 2-phase liquid–liquid system. J. Chem. Soc., Chem. Commun. **1994**, 801–802.
13. Templeton, A.C.; Hostetler, M.J.; Warmoth, E.K.; Chen, S.W.; Hartshorn, C.M.; Krishnamurthy, V.M.; Forbes, M.D.E.; Murray, R.W. Gateway reactions to diverse, polyfunctional monolayer-protected gold clusters. J. Am. Chem. Soc. **1998**, *120*, 4845–4849.
14. Templeton, A.C.; Wuelfing, M.P.; Murray, R.W. Monolayer-protected cluster molecules. Acc. Chem. Res. **2000**, *33*, 27–36.
15. Toshima, N.; Yonezawa, T. Bimetallic nanoparticles—Novel materials for chemical and physical applications. New J. Chem. **1998**, *22*, 1179–1201, and references cited therein.
16. Schmidt, G.; Bäumle, M.; Geerkens, M.; Heim, I.; Osemann, C.; Sawitowski, T. Current and future applications of nanoclusters. Chem. Soc. Rev. **1998**, *28*, 179–185.
17. Clapham, B.; Reger, T.S.; Janda, K.D. Polymer-supported catalysis in synthetic organic chemistry. Tetrahedron **2001**, *57*, 4637–4662.
18. Aiken, J.D.; Finke, R.G. A review of modern transition-metal nanoclusters: Their synthesis, characterization, and applications in catalysis. J. Mol. Catal., A Chem. **1999**, *145*, 1–44.
19. Centi, G.; Perathoner, S. Oxidation catalysts: New trends. Curr. Opin. Solid State Mater. Sci. **1999**, *4*, 74–79.
20. Forster, S.; Antonietti, M. Amphiphilic block copolymers in structure-controlled nanomaterial hybrids. Adv. Mater. **1998**, *10*, 195–217.
21. Klingelhofer, S.; Heitz, W.; Greiner, A.; Oestreich, S.; Forster, A.M. Preparation of palladium colloids in block copolymer micelles and their use for the catalysis of the heck reaction. J. Am. Soc. Chem. **1997**, *119*, 10116–10120.
22. Li, Y.; Hong, X.M.; Collard, D.M.; El-Sayed, M.A. Suzuki cross-coupling reactions catalyzed by palladium nanoparticles in aqueous solution. Org. Lett. **2000**, *2*, 2385–2388.
23. Underhill, R.S.; Liu, G. Preparation and performance of Pd particles encapsulated in block copolymer nanospheres as a hydrogenation catalyst. Chem. Mater. **2000**, *12*, 3633–3641.
24. Crooks, R.M.; Zhao, M.; Sun, L.; Chechik, V.; Yeung, L.K. Dendrimer-encapsulated metal nanoparticles: Synthesis, characterization, and applications to catalysis. Acc. Chem. Res. **2001**, *34*, 181–190.
25. Niu, Y.; Yeung, L.K.; Crooks, R.M. Size-selective hydrogenation of olefins by dendrimer-encapsulated palladium nanoparticles. J. Am. Chem. Soc. **2001**, *123*, 6840–6846.
26. Mecking, S.; Thomann, R.; Holger, F.; Sunder, A. Preparation of catalytically active palladium nanoclusters in compartments of amphiphilic hyperbranched polyglycerols. Macromolecules **2000**, *33*, 3958–3960.
27. Boal, A.K.; Galow, T.H.; Ilhan, F.; Rotello, V.M. Binary and ternary polymer-mediated "bricks and mortar" self-assembly of gold and silica nanoparticles. Adv. Funct. Mater. **2001**, *11*, 461–465.
28. Galow, T.H.; Drechsler, U.; Hanson, J.A.; Rotello, V.M. Highly reactive heterogeneous heck and hydrogenation catalysts constructed through "bottom-up" nanoparticle self-assembly. Chem. Commun. **2002**, 1076–1077.
29. Shipway, A.N.; Katz, E.; Willner, I. Nanoparticle arrays on surfaces for electronic, optical, and sensor applications. ChemPhysChem **2000**, *1*, 18–52.
30. Murray, C.B.; Kagan, C.R.; Bawendi, M.G. Synthesis and characterization of monodisperse nanocrystals and close-packed nanocrystal assemblies. Annu. Rev. Mater. Sci. **2000**, *30*, 545–610.
31. Forster, S.; Plantenberg, T. From self-organizing polymers to nanohybrid biomaterials. Angew. Chem., Int. Ed. Engl. **2002**, *41*, 688–714.
32. Russell, T.P. Copolymers at surfaces and interfaces. Curr. Opin. Colloid Interface Sci. **1996**, *1*, 107–115.
33. Rutnakornpituk, M.; Thompson, M.S.; Harris, L.A.; Farmer, K.E.; Esker, A.R.; Riffle, J.S.; Connolly, J.; St. Pierre, T.G. Formation of cobalt nanoparticle dispersions in the presence of polysiloxane block copolymers. Polymer **2002**, *43*, 2337–2348.
34. Moore, R.B.; Bittencourt, D.; Gauthier, M.; Williams C.E.; Eisenberg, A. Small-angle X-ray scattering investigations of ionomers with variable-length side-chains. Macromolecules **1991**, *24*, 1376–1382.
35. Ng Cheong Chan, Y.; Schrock, R.R.; Cohen, R.E. Synthesis of single silver nanoclusters within spherical microdomains in block copolymer films. J. Am. Chem. Soc. **1992**, *114*, 7295–7296.
36. Fogg, D.E.; Radzilowski, L.H.; Blanski, R.; Schrock, R.R.; Thomas, E.L. Fabrication of quantum dot/polymer composites: Phosphine-functionalized block coploymers as passivating hosts for cadmium selenide nanoclusters. Macromolecules **1997**, *30*, 417–426.
37. Boal, A.K.; Ilhan, F.; DeRouchey, J.E.; Thurn-Albrecht, T.; Russell, T.P.; Rotello, V.M. Self-assembly of nanoparticles into structured spherical

and network aggregates. Nature **2000**, *404*, 746–748.
38. Boal, A.K.; Gray, M.; Ilhan, F.; Clavier, G.M.; Kapitzky, L.; Rotello, V.M. Bricks and mortar self-assembly of nanoparticles. Tetrahedron **2002**, *58*, 765–770.
39. Frankamp, B.L.; Uzun, O.; Ilhan, F.; Boal, A.K.; Rotello, V.M. Recognition-mediated assembly of nanoparticles into micellar structures with diblock copolymers. J. Am. Chem. Soc. **2002**, *124*, 892–893.
40. Frankamp, B.L.; Boal, A.K.; Rotello, V.M. Controlled interparticle spacing through self-assembly of Au nanoparticles and poly(amidoamine) dendrimers. J. Am. Chem. Soc. **2002**, *124*, 15146–15147.
41. Vossmeyer, T.; Guse, B.; Besnard, I.; Bauer, R.E.; Müllen, K.; Yasuda, A. Gold nanoparticle/polyphenylene dendrimer composite films: Preparation and vapor-sensing properties. Adv. Mater. **2002**, *14*, 238–242.
42. Möller, M.; Kunstle, H.; Kunz, M. Inorganic nanoclusters in organic glasses—Novel materials for electro-optical applications. Synth. Met. **1991**, *41*, 1159–1162.
43. Spatz, J.P.; Sheiko, A.; Möller, M. Ion-stabilized block copolymer micelles: Film formation and intermicellar interaction. Macromolecules **1996**, *29*, 3220–3226.
44. Spatz, J.P.; Roescher, A.; Möller, M. Gold nanoparticles in micellar poly(styrene)-*b*-poly(ethylene oxide) films—Size and interparticle distance control in monoparticulate films. Adv. Mater. **1996**, *8*, 337–340.
45. Spatz, J.P.; Mössmer, S.; Hartmann, C.; Möller, M.; Herzog, T.; Krieger, M.; Boyen, H.-G.; Ziemann, P.; Kabius, B. Ordered deposition of inorganic clusters from micellar block copolymer films. Langmuir **2000**, *16*, 407–415.
46. Decher, G. Fuzzy nanoassemblies: Towards layered polymeric multicomposites. Science **1997**, *277*, 1232–1237.
47. Schmidt, G.; Bäumle, M.; Beyer, N. Ordered two-dimensional monolayers of Au_{55} clusters. Angew. Chem. Int. Ed. **2000**, *39*, 181–183.
48. Hao, E.; Lian, T. Buildup of polymer/Au nanoparticle multilayer thin films based on hydrogen bonding. Chem. Mater. **2000**, *12*, 3392–3396.
49. Schmitt, J.; Decher, G.; Dressick, W.J.; Brandow, S.L.; Geer, R.E.; Shashidhar, R.; Calvert, J.M. Metal nanoparticle/polymer superlattice films: Fabrication and control of layer structure. Adv. Mater. **1997**, *9*, 61–65.
50. Feldheim, D.L.; Grabar, K.C.; Natan, M.J.; Mallouk, T.E. Electron transfer in self-assembled inorganic polyelectrolyte/metal nanoparticle heterostructures. J. Am. Chem. Soc. **1996**, *118*, 7640–7641.
51. Freeman, R.G.; Grabar, K.C.; Allison, K.J.; Bright, R.M.; Davis, J.A.; Guthrie, A.P.; Hommer, M.B.; Jackson, M.A.; Smith, P.C.; Walter, D.G.; Natan, M.J. Self-assembled metal colloid monolayers: An approach to SERS substrates. Science **1995**, *267*, 1629–1632.
52. Gao, M.; Zhang, X.; Yang, B.; Li, F.; Shen, J. Assembly of modified CdS particles/cationic polymer based on electrostatic interactions. Thin Solid Films **1996**, *284–285*, 242–245.
53. Sun, S.; Anders, S.; Haumann, H.F.; Thiele, J.-U.; Baglin, J.E.E.; Thomson, T.; Fullerton, E.E.; Murray, C.B.; Terris, B.D. Polymer mediated self-assembly of magnetic nanoparticles. J. Am. Chem. Soc. **2002**, *124*, 2884–2885.
54. Ng Cheong Chan, Y.; Schrock, R.R.; Cohen, R.E. Synthesis of silver and gold nanoclusters within microphase-separated diblock copolymers. Chem. Mater. **1992**, *4*, 24–27.
55. Cummins, C.C.; Schrock, R.R.; Cohen, R.E. Synthesis of ZnS and CdS within ROMP block copolymer microdomains. Chem. Mater. **1992**, *4*, 27–30.
56. Lopes, W.A.; Jaeger, H.M. Hierarchical self-assembly of metal nanostructures on diblock copolymer scaffolds. Nature **2001**, *414*, 735–738.
57. Lopes, W.A. Nonequilibrium self-assembly of metals on diblock copolymer templates. Phys. Rev., E **2002**, *65*, 031606 (1–14).

Polymer–Nanoparticle Composites

Kevin Sill
Seunghoo Yoo
Todd Emrick
University of Massachusetts, Amherst, Massachusetts, U.S.A.

INTRODUCTION

The rapid development in the chemistry of nanoparticles over the past 20 years has driven tremendous advances in the field. While there remains significant interest in the use of nanoparticles as fillers in polymer materials to enhance physical and mechanical properties, many are now engaging in research efforts that focus on precise structures of nanoparticles in polymers, including their assembly in arrays and along interfacial boundaries. The combination of precise organic chemistry on nanoparticle materials has led to developments in the growth of polymers from nanoparticle surfaces, allowing one to tailor the properties of the particles by the choice of polymer and its inherent functionality. Such nanoparticle–polymer composite materials are generating interest in several applications, including electronic and optical materials, based on the properties of metallic and semiconductor particles used. In addition, efforts in catalysis are under way using nanoparticles in polymer scaffolds. These and other topics form the basis of this review article, which will start in the next section with a brief account of early studies in the field, and progress from clay-based composites to metallic- and semiconductor-based nanoparticle–polymer composites.

NANOPARTICLES IN POLYMERS: FROM CLAY TO SEMICONDUCTORS

The integration of nanoparticles into polymers is of significant theoretical and experimental interest, with a rich history in the polymer and engineering communities. Inorganic fillers have been used for some time in conjunction with organic polymer materials, largely in an effort to enhance the physical and mechanical properties over those of the polymers alone. In the mid-19th century, research efforts by Charles and Nelson Goodyear, pioneers in the chemistry of rubber, showed that vulcanized rubber can be toughened significantly by the addition of zinc oxide and magnesium sulfate.[1–3] In the early 1900s, Leo Baekeland investigated the use of silicate clay in phenolic resins that helped Bakelite™, the first mass-produced synthetic polymer composite, become a commercial success.[1–4] Many additional commodity materials have been developed based on the enhanced properties that result from filling rubber with clay.[5] More recently, it was found that rubber particles embedded in nylon and other polymer matrices afford composites with outstanding impact resistance.[6] These are among the many examples of composite materials on which new research efforts have been constructed. These efforts have focused on the integration of a variety of nanoscopic materials into polymers and the impact of these composites over a broad spectrum of applications, from engineering plastics to electronic materials to biotechnology.

There is no question that nanotechnology is revolutionizing the scientific approach across many disciplines, including chemistry, biology, materials science, engineering, and theory.[7–10] For example, research centered on nanoscopic materials extends from the semiconductor industry, where the ability to produce nanometer-scale features leads to faster and less expensive transistors,[11–13] to biotechnology, where, e.g., luminescent nanoparticles are extremely interesting as bioprobes.[14–16] Tremendous advances in characterization technology is a common thread through all disciplines in nanotechnology, where nanoscopic objects and materials can be visualized with greatly improved resolution relative to only 10–15 years ago. Fig. 1 illustrates a few examples of ''nanomaterials'' with dimensions of 1 micron and less. In the nanoparticle–polymer area, current efforts now reach far beyond (or below) the use of micron-sized particle fillers (e.g., layered silicates) and into very small metallic and semiconductor nanoparticles as small as 1–2 nm in diameter.

This article will highlight recent research in the area of polymer–nanoparticle composites, focusing especially on preparative techniques that afford well-dispersed nanocomposites, because such dispersion is a major target in the field. A leading discussion of clay-based nanocomposites will be followed by a focus on metallic and semiconductor-based hybrid materials. The use of polymers as a means to provide exquisite order to nanoparticles will be described, based on the ability of polymer

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013728

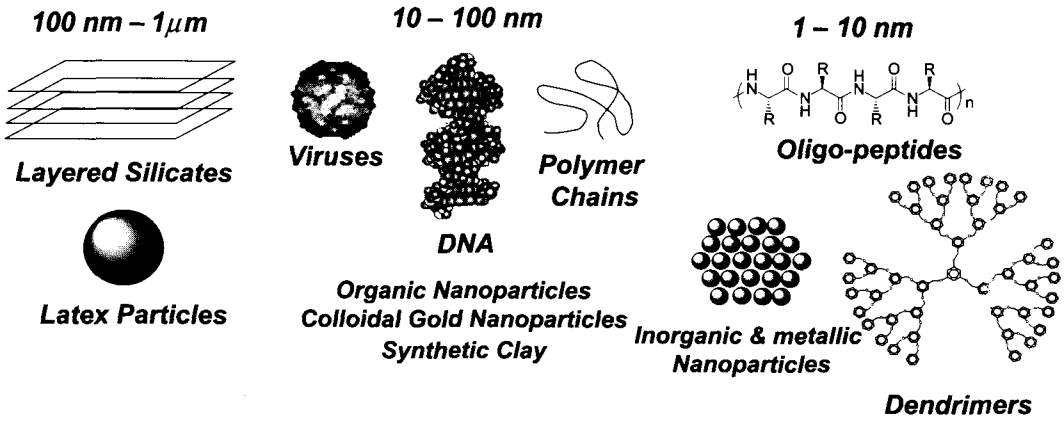


Fig. 1 Building blocks for nanotechnology. (*View this art in color at www.dekker.com.*)

materials to assemble into nanostructures. Finally, potential applications of polymer–nanoparticle composites will be discussed, with a special focus on the use of dendritic polymers and nanoparticles for catalysis.

Clay–Nanoparticle Polymer Composites

The use of clay (i.e., hydrated silicates of aluminum) in combination with polymer materials can provide significant advantages in physical properties relative to the polymers alone.[17] Early discoveries in this area described blends of layered silicate nanoparticles with nylon-6 to afford composites with greatly enhanced thermal and mechanical properties.[18–20] Further advances in such composites in terms of melt processing allow these materials to be prepared in the absence of organic solvents.[21] Manias and coworkers have developed polypropylene–montmorillonite composites by varying the volume fraction of montmorillonite nanoparticles and varying block lengths of polypropylene–styrene or polypropylene–poly(methylmethacrylate) copolymers to probe mechanical and thermal properties.[22] For example, a 3% inorganic loading into polypropylene gave a 30% increase in Young's modulus and a 30°C increase in heat deflection temperature when compared to native polypropylene. In addition, there is significant interest in the fire-retardant properties of polymer–clay nanocomposites, as integration of clay into the structure inhibits the inherent flammability of most organic-based polymers.[23]

Significant challenges are associated with blending polymers and nanoparticles to afford homogeneous, well-dispersed inorganic material within the polymer. In order for dispersion to be achieved, the entropic penalty associated with addition of the nanoparticles must be balanced by favorable enthalpic interactions. Typical polymer–clay hybrids composed of layered nanoparticles (silicates, talc, and mica) are aggregated to some degree, as the immiscibility of clay in the polymer leads to a very close proximity of sheets to one another.[24] The degrees of dispersion in these composites are generally referred to as unmixed (highly aggregated), intercalated (minimally aggregated), and exfoliated (well dispersed). In the intercalated case, polymer chains interpenetrate stacked silicate layers with small separation distances (few nanometers) between the layers. In the exfoliated or delaminated morphology, the silicate layers are well dispersed within the polymer matrix (Fig. 2). When polar polymers are used, exfoliation can be achieved by the addition of a surfactant to the material, typically a long-chain alkylammonium salt. However, for nonpolar

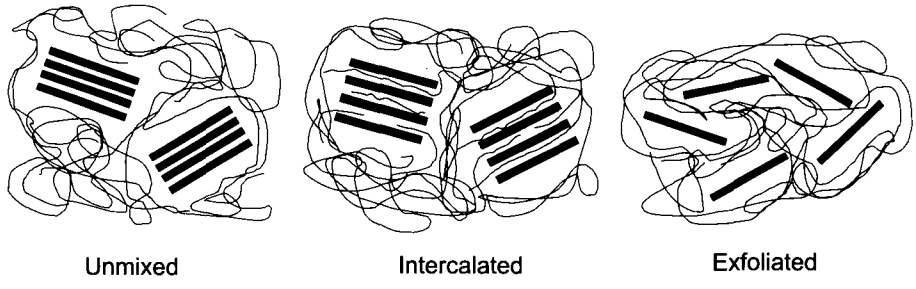


Fig. 2 Depiction of three types of clay–polymer hybrid materials showing different levels of particle dispersion.

polymers such as poly(ethylene) and poly(propylene), the addition of a surfactant is not sufficient to overcome the entropic penalty; thus a functional comonomer such as methyl methacrylate must be incorporated into the nonpolar polymer to allow nanoparticle dispersion within the matrix.[22] Advances in processing have also led to decreased aggregation in clay–polyethylene materials, such as the use of supercritical CO_2 during polyethylene extrusion.[25,26] Because the physical properties of these composites depend on the ability to produce controlled intercalation or exfoliation, the interest in small molecule and polymeric additives tuned to this target is growing significantly.

Recent studies by Balazs and coworkers describe models based on self-consistent field theory to determine the interactions needed to allow polymer chains to penetrate nanoparticle layers or sheets, leading to a further increase in the composite's physical properties.[27] Their findings indicate that the introduction of small amounts of functionalized polymers to the homopolymer system will greatly increase the thermodynamic stability of the clay–polymer composite and increase the degree of intercalation or exfoliation in the system. Based on these and other findings, a much research has been dedicated to optimizing the properties of these clay–polymer systems based on blend ratio, degree of cross-linking, curing behavior, and size and dispersity of the nanoparticles.[28]

Composites of Polymers with Metallic and Semiconductor Nanoparticles

Related to research efforts in clay–nanoparticle composite materials are efforts to integrate metallic and semiconductor nanoparticles into polymer materials. While targets such as controlled dispersion and self-assembly parallel those of the clay-based materials, the metal and semiconductor-based composites offer clearly distinct materials in that electronic, magnetic, and optical properties offered by the particles are now present. For example, nanoparticles composed of gold, cobalt, palladium, copper, and cadmium selenide are of great interest for their potential role in nanotechnology applications.[29–34] Nanoparticles such as these are generally synthesized by "bottom-up" chemistry that affords nanoparticles with narrow size distributions. Very small nanoparticles (i.e., <10 nm) are especially interesting for the quantum confinement effects inherent in particles of that size.[29,30,35] Semiconductor nanoparticles such as cadmium selenide and cadmium telluride are of special interest because of their narrow (ca. 20–25 nm FWHM) and tunable (based on size) fluorescence emission wavelengths.[30,36] Key to the behavior of these nanoparticles is their organic ligand shell that provides a protective encapsulating layer to prevent aggregation and oxidative or chemical degradation, all of which serve to diminish or destroy the nanoscopic integrity and properties of the nanoparticles (Fig. 3).

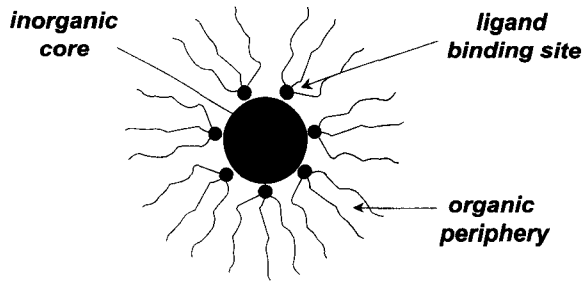


Fig. 3 Representation of nanoparticles passivated by organic ligands.

To provide dispersion of these organic-capped nanoparticles in a polymer matrix, there must be favorable interactions between the nanoparticle ligands and the polymer environment. Several approaches have been investigated, such as 1) mimicking the nanoparticle ligand coverage with side chains on the polymer structure; 2) preparing polymers that contain ligands to bind to the nanoparticles; 3) the attachment of polymers to the nanoparticles by functionalized chain-ends, also known as "grafting-to"; and 4) the attachment of polymerization initiators to nanoparticle surfaces, followed by polymer growth from outward from the surface, also known as "grafting-from." Method 4 is proving to be an excellent approach whereby a high density of initiators can be attached to nanoparticles, and a diverse range of monomers can then be grown off the particles; this will serve as the focus of the next section.

POLYMER GROWTH FROM NANOPARTICLE SURFACES

Functionalization of nanoparticles with polymer chains opens new avenues in nanostructured materials and composites by tailoring the interactions of the nanoparticle with its environment based on the properties of the polymer encapsulant. This concept may be used to introduce an enormous diversity of chemical functionality into composite systems that can be used directly, or modified further. This chemical functionality can be tailored to provide specific solubility or miscibility, or used for further coupling chemistry, e.g., to reactive surfaces.

There are two general methods used for attachment of polymers to nanoparticle surfaces. The first is coupling of end-functionalized polymers onto the surface of

nanoparticles in a "grafting-to" approach. Although this procedure is commonly used because of the simplicity of the approach, less than optimum grafting density may result, because of the steric shielding that arises on placement of each successive polymer chain onto the nanoparticle. Some recent examples using this "grafting-to" method include thiol-functionalized polystyrene (PS) grafted to gold nanoparticles,[37] pyridine-functionalized poly(ethylene glycol) grafted to CdSe nanoparticles,[38] thiol-functionalized poly(caprolactone) grafted to CdS nanoparticles,[39] and a variety of copolymers with end-thiol functionality, including water-soluble particles, for stabilization of gold nanoparticles.[40]

The second method used to connect polymers and nanoparticles involves polymerization from the nanoparticle surface, using nanoparticles functionalized appropriately with polymerization-active ligands. The appeal of this "grafting-from" method is derived from the facile attachment of small molecules to the nanoparticle surface relative to chain-end functionalized polymers, and the ability to subsequently grow a compact shell of polymers outward from the surface. Critical to this "grafting-from" process is the compatibility of the nanoparticle with the polymerization conditions chosen, such that neither the attachment of functional ligands nor the polymerization process appreciably alters the fundamental features of the nanoparticles. Although these requirements pose a considerable challenge, much progress has been made in recent years. It should be noted that nanoparticle surfaces differ from flat surfaces as substrates in the "grafting-from" process, as higher surface curvature of the spherical nanoparticle can reduce steric crowding during chain growth, which may lead to more effective polymerization from the nanoparticles.

Polymerization techniques utilized in the "grafting-from" method include controlled radical polymerization, "living" anionic polymerization, and ring-opening metathesis polymerization (ROMP).[41] Optimum control over the properties of these composites can be achieved by varying the type of nanoparticle and by adjusting the composition, molecular weight, and functionality of the attached polymer shell. ROMP is proving to be especially useful in grafting from nanoparticles, due in part to the excellent chemical tolerance of the ruthenium-based catalysts used in ROMP, and the mild conditions under which they perform. Watson and coworkers recently reported the surface functionalization of gold nanoparticles with polymerizable moieties that were used to prepare nanocomposites by ROMP.[41] This was accomplished by functionalization of gold nanoparticles with 1-mercapto-10-(*exo*-5-norbornen-2-oxy)decane, where the thiol group binds to the gold surface, and a ROMP-active

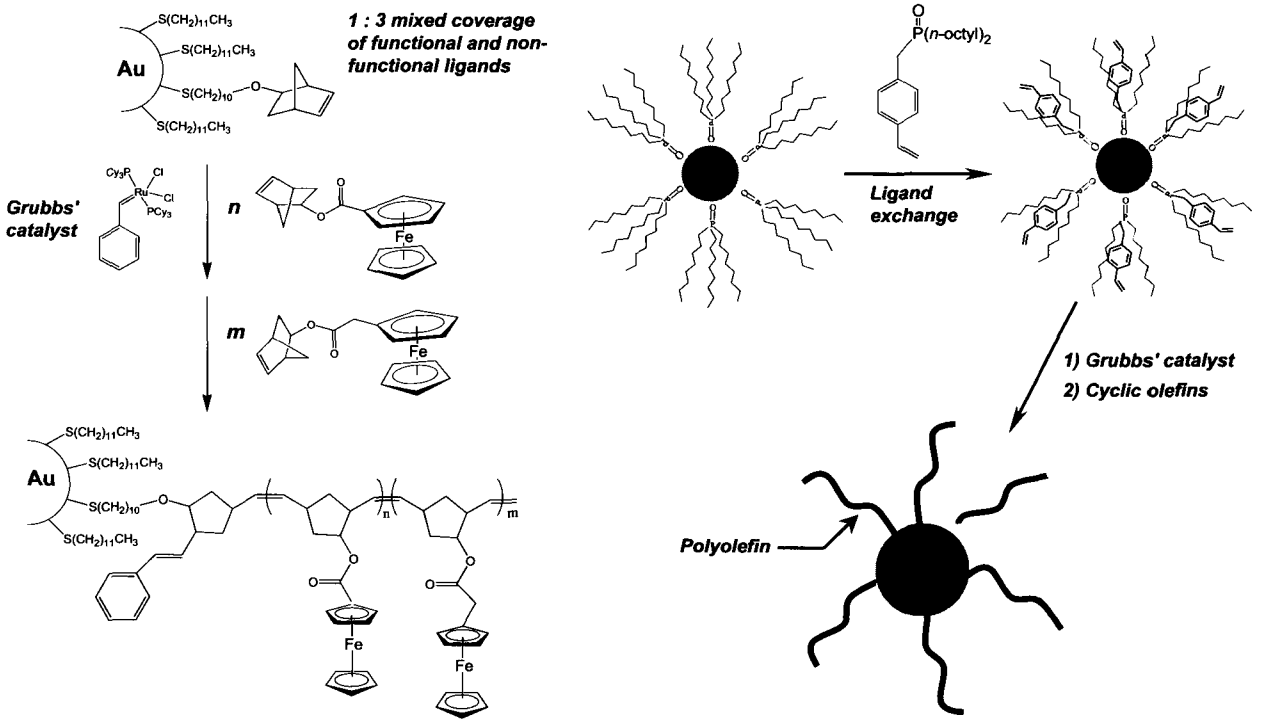


Fig. 4 Schematic representation of the preparation of polyolefin-functionalized gold (left) and CdSe (right) nanoparticles by ring-opening metathesis techniques. (*View this art in color at www.dekker.com.*)

functionality is left available for polymerization. *bis*(Tricyclohexylphosphine)benzylidine ruthenium dichloride (i.e., first-generation Grubbs catalyst) was used as a surface-immobilized ROMP catalyst, and a ferrocene-substituted norbornene was then polymerized from the nanoparticle surface (Fig. 4). The "living" nature of this ROMP method was demonstrated by the preparation of block copolymers from the gold surface, using two different ferrocene esters. Cyclic voltammetry of the composite exhibited two distinguishable waves characteristic of oxidation/reduction of the different ferrocenyl moieties. From the integrated current associated with these waves, a copolymer composition of 1.4:1 [poly(*exo*-5-norbornen-2-yl-ferrocenecarboxylate):poly(*exo*-5-norbornen-2-yl-ferroceneacetate)] was calculated.

Recently, Skaff and coworkers applied the "grafting-from" concept to semiconductor nanoparticles, using ca. 3-nm CdSe nanoparticles and cyclooctene-based monomers.[42] This involved the synthesis and use of a new ligand for CdSe, *p*-vinylbenzyl-di-*n*-octylphosphine oxide, prepared from di-*n*-octylphosphine oxide and 4-vinylbenzyl chloride. This ligand is first used to functionalize the nanoparticles, followed by metathesis chemistry of the vinyl moiety with Grubbs' catalyst (generation 1) to afford a macroinitiator that is suitable for polymerization of cyclic olefins (Fig. 4). For example, cyclooctene was polymerized outward from the surface of the macroinitiator nanoparticle, and excellent particle dispersion was observed in the resulting poly(cyclooctene)–CdSe hybrid material. In addition, other cyclic olefins such as *exo*-7-oxanorbornene-2,3-dicarboxylic anhydride and *exo*-*N*-ethyl-7-oxanorbornene-2,3-dicarboxylimide proved applicable to this approach, and it is expected that many other cyclic olefins could be polymerized from this versatile nanoparticle-based macroinitiator.

Jordan and coworkers used ω-functionalized self-assembled monolayers of thiolates on gold nanoparticles to initiate living cationic ring-opening polymerizations of 2-oxazoline monomers.[43] *N*,*N*-di-*n*-Octadecylamine was introduced as a terminating agent to provide a hydrophobic region and thus afford an amphiphilic polymer shell around the gold core. A schematic illustration of the preparation of such nanoparticles is given in Fig. 5. Tethered polymers were liberated from the nanoparticles by using NaCN solution and analyzed by matrix-assisted laser desorption ionization time-of-flight (MALDI TOF) mass spectrometry. These polymerizations were observed to proceed in a living fashion, as judged from the linear relationship between molecular weight of the grafted polymer after cleavage from the nanoparticle surface and reaction time.

Surface-initiated anionic polymerization from nanoparticle surfaces has also been reported, e.g., using 1,1-diphenylethylene with a chlorosilane end group to functionalize silica nanoparticles, followed by anionic polymerization of styrene (Fig. 6).[44] Although the molecular weight distributions ($M_w/M_n = 1.2–1.8$) of the surface-bound polymers were broader than typical solution anionic polymerization, this is a rare example where living anionic polymerization is performed on nanoparticle surfaces; it will be interesting to see the progress in this area in terms of which nanoparticle compositions can be used in conjunction with the living anionic chemistry.

Atom transfer radical polymerization (ATRP) is among the most extensively exploited methods for polymer

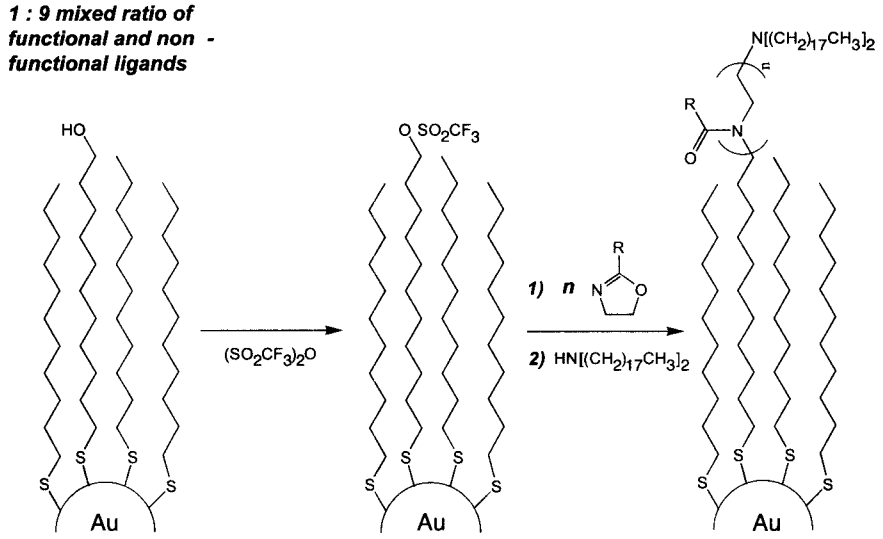


Fig. 5 Reaction scheme of gold-nanoparticle-initiated cationic polymerization of 2-substituted oxazolines.

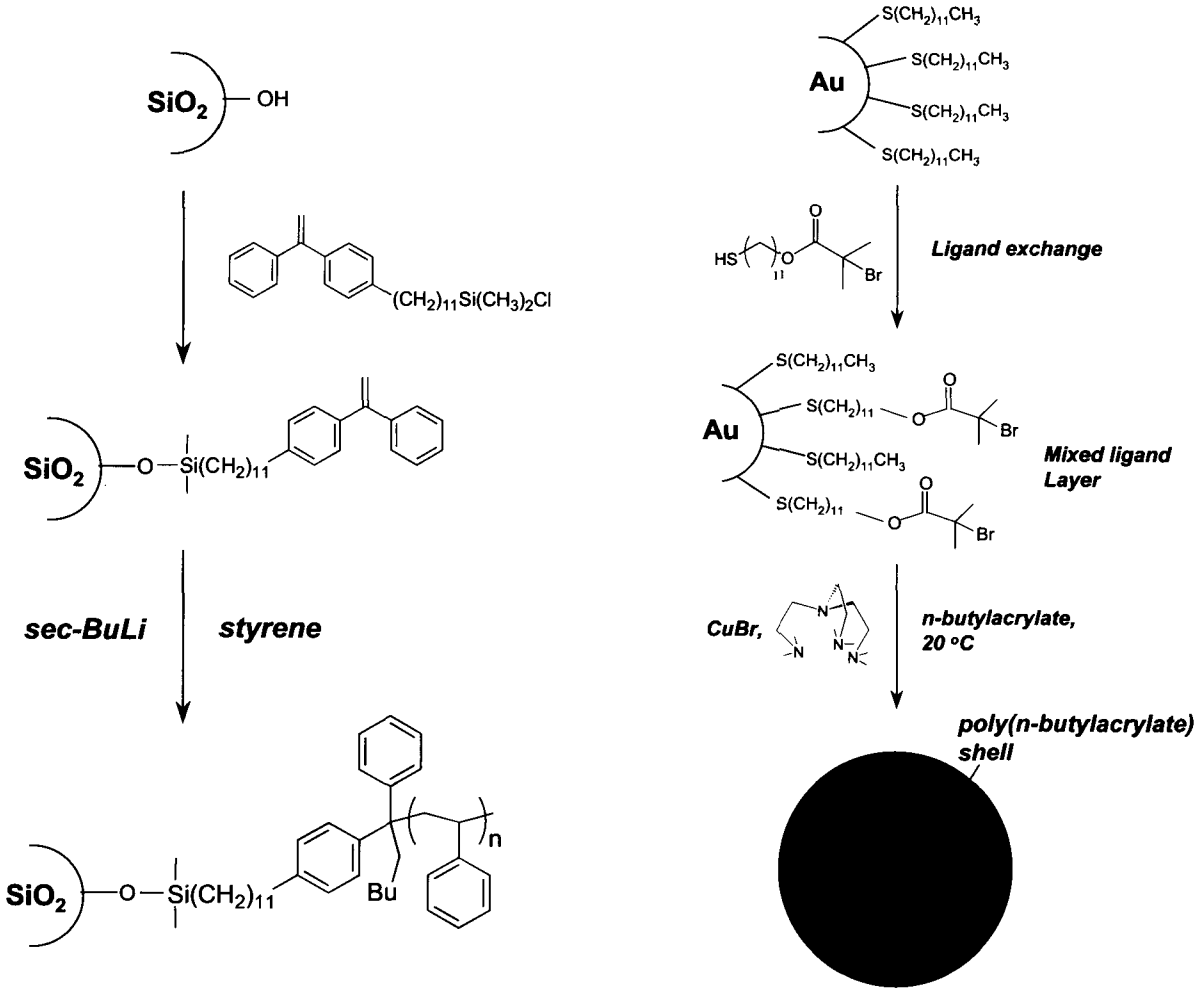


Fig. 6 Schematic representation of polymer growth from nanoparticle surfaces by anionic polymerization (left) and atom transfer radical polymerization (right). (*View this art in color at www.dekker.com.*)

growth from nanoparticle surfaces. Recent studies on ATRP from silica[45–47] and core–shell CdS/SiO_2 nanoparticles[48] have demonstrated the ability to grow well-defined PS or poly(methylmethacrylate) (PMMA) layers from these modified inorganic surfaces. In these studies, some interesting effects of nanoparticle size were reported. In the case of 75-nm-diameter silica nanoparticles, good molecular weight control of PS was observed. However, the growth-from method on larger silica nanoparticles (300-nm diameter) did not show characteristics of a living polymerization. This was attributed to the high ratio of monomer to initiator per unit mass of silica used in the latter case. However, in each case the CdS/SiO_2 nanoparticles were dispersed throughout the polymer, with no signs of aggregation.

Other examples of the preparation of polymer-attached particles by "grafting-from" ATRP methods were reported by Hallensleben and coworkers.[49,50] In this case, *tris*(2-dimethylaminoethyl)amine was used as the ATRP ligand, which enabled the polymerization to proceed at room temperature (Fig. 6). This low reaction temperature has the advantage of suppressing side reactions, such as desorption of thiol-functionalized ATRP initiators from the gold surface, and thermal initiation of *n*-butylacrylate monomer.[50] Gold nanoparticles had varying numbers of the polymer chains, which resulted in a relatively broad molecular weight distribution, as judged by the gel permeation chromatography (GPC) trace of poly(*n*-butylacrylate)-grafted gold nanoparticles.

An interesting approach to prepare silica nanoparticle–polymer hybrid materials has been demonstrated by Mori et al., where a hyperbranched polymer (HBP) was grown from the particle surface.[51] Typical hyperbranched polymerizations that proceed by polycondensation chemistry would result in cross-linked networks when performed in the presence of functionalized nanoparticles that participate in the chemistry. However, in this case self-condensing vinyl polymerization[52] of an acrylic

"inimer" was used, where the inimer functions as both a monomer and an initiator. This inimer is designated as "AB*," where A is a vinyl group and B* is a functional group that can be transformed to an active center capable of initiating polymerization. Bulk polymerization of the inimer, 2-(2-bromopropionyloxy)ethyl acrylate, with the silica nanoparticles functionalized with 2-bromoisobutyryl fragment (B*) was performed via ATRP, resulting in the formation of HBP shell on the silica surface (Fig. 7). These efforts are promising for providing diverse polymer architecture to nanoparticle-based composites, because HBPs have distinctive chemical and physical properties due to their globular shape and many end groups. This may also provide a rapid alternative to the synthesis of dendrimer–nanoparticle hybrid materials, which require the tedious dendritic synthesis.[53,54]

ORDERED ASSEMBLIES OF NANOPARTICLE–POLYMER COMPOSITES

Well-ordered nanoparticles within polymers are expected to provide access to new materials that combine the unique physical properties of the particles with the superior processibility of the polymers. The ability to control the assembly of nanoparticles in materials promises advances in information storage, nanoelectronics, and quantum computing. A variety of nanoparticle assembly methods are under investigation, e.g., into superlattice structures reported by Murray and Bawendi for CdSe nanoparticles with aliphatic ligands of various size.[33,55] These highly ordered structures are prepared by slow crystallization to give ordered nanoparticle assemblies with remarkably few defects.

Theoretical models of Balasz and coworkers are addressing the impact of nanoparticles on phase separation in A–B diblock copolymers.[32,56–58] These calculations are driven by a combination of self-consistent field theory, which is preferred for diblock copolymers, and density functional theory, which is optimal for studying ordered colloidal particles. The theory accounts for key parameters such as nanoparticle size, chain length of the polymer blocks, and the relevant interaction parameters, including those between nanoparticles and polymer A, nanoparticles and polymer B, and polymer A with polymer B. These models show that nanoparticles can be directed to the center or the edges of phase-separated lamella or spherical domains by choosing systems with

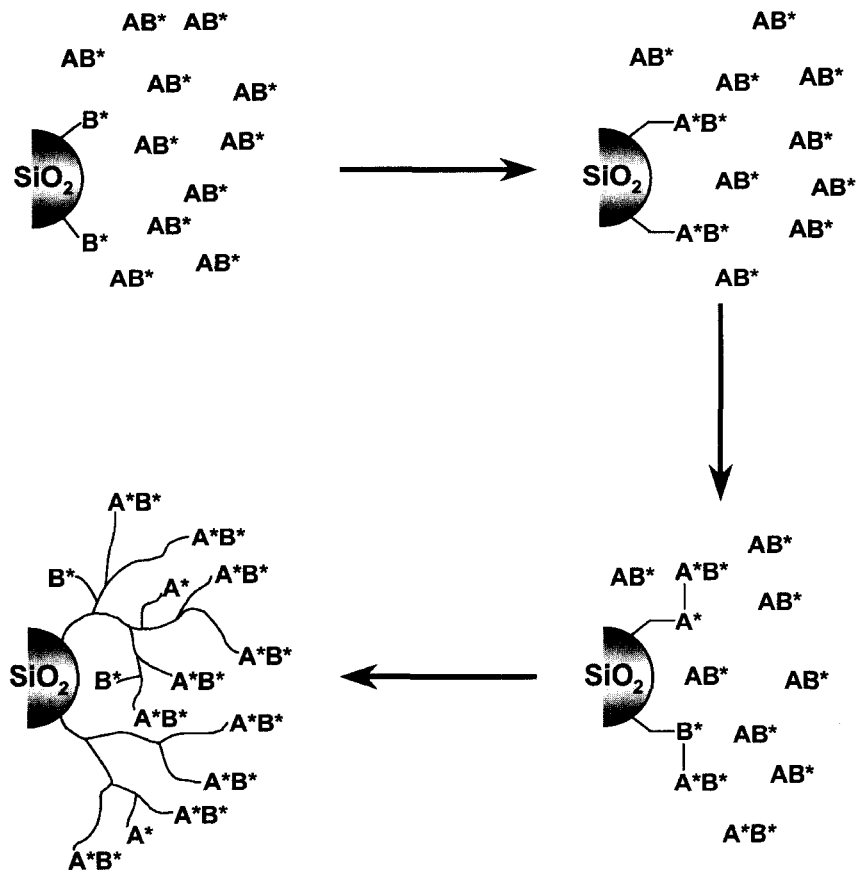


Fig. 7 Schematic representation of self-condensing vinyl polymerization of AB* inimer from functionalized silica nanoparticles.

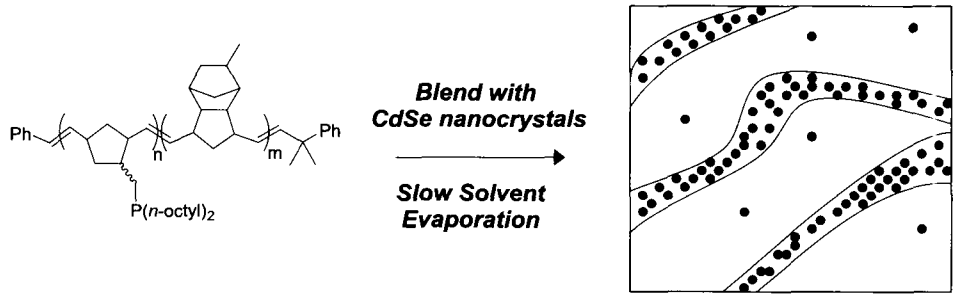


Fig. 8 Schematic representation of nanoparticle segregation to the ligand-rich phase of an AB diblock copolymer.

the appropriate parameters. The increased complexity of the system that results from placing ligands on the nanoparticles would certainly have an impact on the outcome of these calculations.

A report by Fogg et al. in 1997 on nanoparticle–polymer assemblies was one of the early papers to address specific polymer–nanoparticle interactions to promote assembly in phase-separated systems.[59] In this study, a diblock copolymer was prepared by ROMP, where one of the blocks contained phosphine or phosphine oxide moieties as nanoparticle ligands. This diblock copolymer was annealed in the presence of CdSe nanoparticles, resulting in polymeric phase separation and the nearly exclusive presence of the nanoparticles within the functionalized block, as shown by transmission electron microscopy (TEM) analysis, and illustrated schematically in Fig. 8.

Several research groups have recently been active in the field of polymer–nanoparticle composites and assemblies. Rotello and coworkers have been active in self-assembly of gold nanoparticles that are functionalized with ligands that contain recognition sites.[60,61] Specifically, gold nanoparticles were synthesized with aliphatic thiols containing a specific hydrogen bonding sequence, specifically triazine and diaminopyridine derivatives. When these nanoparticles are blended with a polymer or block copolymer containing the complementary hydrogen bond sequence recognition of the two components drives the self-assembly process. Several other methods have been reported for producing ordered assemblies of polymers and nanoparticles. For example, Kotov and coworkers have prepared "nanorainbows" using a multilayer approach,[60–62] where a thin polymer film of poly(diallyldimethylammonium chloride) containing cadmium telluride nanoparticles is deposited onto a quartz slide. Subsequent deposition of larger cadmium telluride nanoparticles with progressively longer wavelengths (e.g., yellow, orange, and red) were then deposited in layer-by-layer fashion. This resulted in a multilayer film that has a color gradient oriented normal to the surface, a material of interest in photonic and electronic devices.

Recent work by Lopes showed self-assembly of various nanoparticles (e.g., gold, silver, indium, and lead) within a PS–PMMA diblock copolymer matrix under nonequilibrium conditions.[63] Under appropriate annealing conditions, the polymer matrix will phase separate into an ordered lamella structure with the nanoparticles residing exclusively in the polystyrene phase because of the more favorable gold–polystyrene interactions. At loadings of 30% gold particles, the polystyrene blocks are capable of conducting an electric charge, essentially converting the polystyrene segments into nanowires. Yuan and coworkers have studied amphiphilic triblock copolymers, such as poly(ethylene oxide-*b*-propylene oxide-*b*-ethylene oxide) as scaffolds for nanostructured silica.[64] When cast from an acidic ethanol solution containing tetraethoxysilane (TEOS), a silica precursor, the copolymer undergoes phase separation to afford an assembly of polymer cylinders. The hydrophilic poly(ethylene oxide) and TEOS form the periphery, and the hydrophobic poly(propylene oxide) is found at the center of the cylinders. The nanostructured silica is formed by calcination of the TEOS, yielding a hollow silica cylinder with a wall thickness of ca. 4 nm.

A different approach to nanoparticle assemblies was reported by Misner and coworkers using templates prepared from diblock copolymers.[65] These templates consist of a cylindrical PMMA phase within a PS matrix, where the cylinder diameter is of the order of 15–20 nm.[66,67] CdSe nanoparticles were driven into these cylinders by capillary forces present on withdrawal of the templates from dilute solutions of CdSe nanocrystals (Fig. 9). This method is quite promising in that nanoparticle coverage in the template was shown to be approximately 70% by TEM (i.e., 70% of the cylinders are filled with nanoparticles), and conditions were developed to the point where >95% of the nanoparticles reside in the cylinders rather than on the PS matrix. Solid-state fluorescence measurements of these templates showed

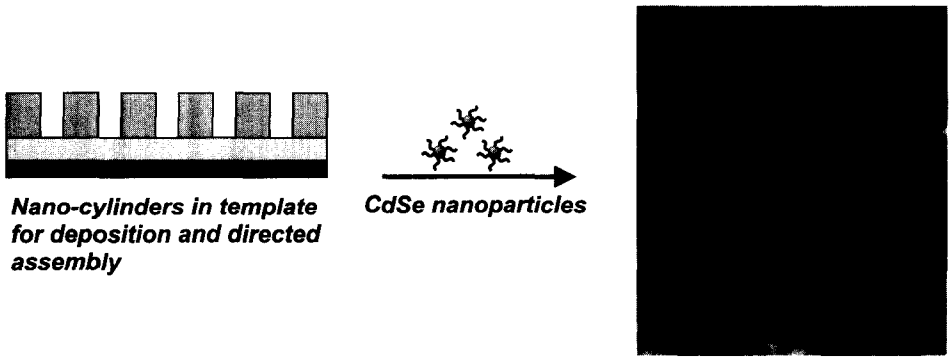


Fig. 9 Nanoparticle-filled templates prepared from self-assembled diblock copolymers are used to direct the assembly of nanoparticles. (*View this art in color at www.dekker.com.*)

that the nanoparticles retained their fluorescence throughout the assembly and substrate removal (acid etching) process, and that fluorescence intensity could be controlled by changing the concentration of solution into which the templates were dipped. New deposition methods currently under investigation are expected to produce templates with complete coverage.[68]

A novel technique for controlling the location of nanoparticles has been demonstrated by Firestone and coworkers.[69] This report shows that by controlling the surface characteristics of silver nanoparticles, one can control their location within a multitier phospholipid bilayer system. Particles that are passivated by lithium dodecyl sulfate reside entirely in the aqueous layer, while alkyl derivatized particles are in the hydrocarbon portion of a lipid bilayer. The use of underivatized silver particles results in assembly at the lipid–water interface (Fig. 10). The authors suggest that the particle interacts with the poly(ethylene oxide) (PEO) chains that have been grafted to some of the surfactant molecules. These hydrophilic polymer chains extend from the bilayer into the aqueous phase, and passivate the bare silver particles. Interestingly,

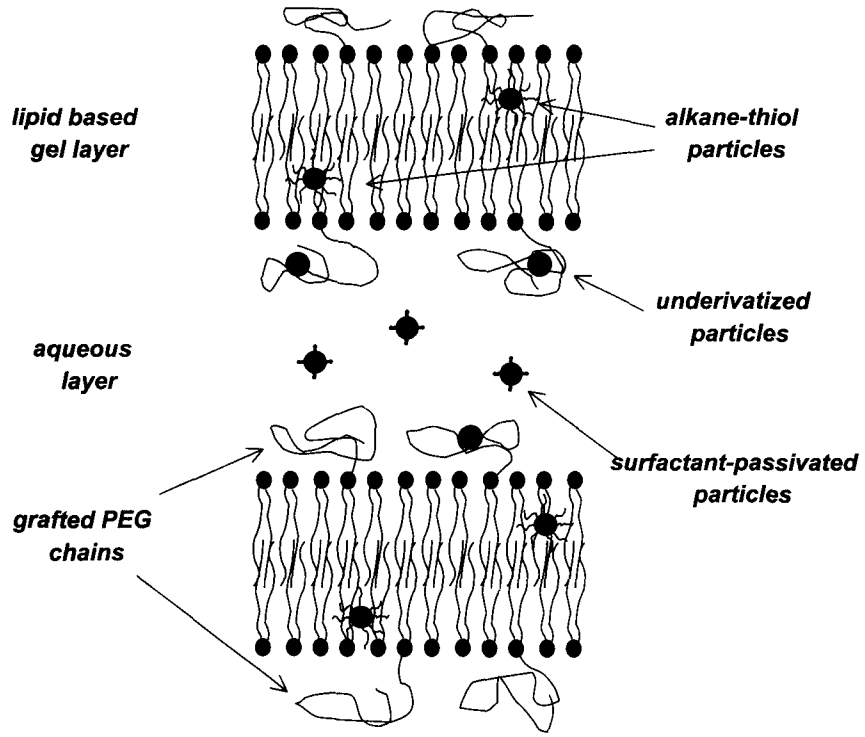


Fig. 10 Schematic representation of a lipid-based gel with selective nanoparticle placement that depends on the nanoparticle ligand environment.

it has been shown by small-angle X-ray scattering that these particles do not aggregate within the fluid, leaving their unique optical and electronic properties intact.

NANOPARTICLE–POLYMER COMPOSITES FOR CATALYSIS

Catalytic methods are of vital importance in the chemical and petroleum industries. Catalysts represent the oldest commercial application of nanoparticles, typically consisting of small metal particles, often in the nanometer size regime, that are dispersed on a microporous oxide. Many studies on the catalytic properties of metal surfaces, such as platinum,[70,71] nickel,[72] iron,[73,74] and molybdenum[75,76] have been explored since the 1920s.[77]

Catalytic processes are greatly affected by the complexity of catalyst surface composition and structure. The catalytic efficiencies of metals increase dramatically with increasing surface-to-volume ratio; thus nanoparticles have inherent advantages. As the high surface energy associated with nanoparticles leads to aggregation, the use of surface-confined stabilizers has been induced. Ideally, these stabilizers should possess rather weak surface activity, as strong adsorption would diminish significantly the catalytic activity. Protective stabilizers can influence not only the stability of the catalyst but also the catalytic properties by creating specific environments around the active sites.

Aside from the conventional commercialized catalytic processes, there have been some very interesting recent studies on nanoparticles in catalysis of organic reactions. For example, El-Sayed reported that Pd nanoparticles stabilized by poly(*N*-vinyl-2-pyrrolidone) (PVP) efficiently catalyze Suzuki cross-coupling reactions in aqueous medium.[78] Successful coupling of arylboronic acids to aryl iodides gave reasonably high yields even though the Pd particles precipitated during the reaction. The reaction rate was observed to depend on the Pd concentration in a linear fashion, indicating that catalysis occurs on the Pd surface. Hydroxyl-terminated poly(amido amine) (PAMAM) dendrimers and polystyrene-*b*-poly(sodium acrylate) were also reported to be efficient stabilizers for the Pd nanoparticles in the Suzuki reactions between phenylboronic acid and iodobenzene.[79] Overall, the careful choice of the type of stabilizer, metal precursors, as well as reaction conditions can lead to the preparation of the desired catalyst system.

Dendrimers have been recognized as suitable hosts for metal nanoparticles for a variety of reasons. Their uniform composition and three-dimensional structure provides a homogeneous environment for the particles, whereas their periphery of terminal end groups can be controlled in terms of functionality to control solubility and allow for catalyst recovery. Substrate selectivity can be affected by the size and polarity of the terminal groups, and the size of the host dendrimer can be varied systematically through well-defined growth chemistry that affords dendritic macromolecules with precise structures and molecular weights.

Dendrimers have been used as both templates and stabilizers for nanoparticles, and have been found to effectively impede their aggregation. In some cases, nanoparticles are confined within the host dendrimer by steric effects, without significant surface passivation. Metal nanoparticles can be sequestered within dendrimers by adsorbing suitable precursor metal ions into the template dendrimer and subsequently reducing the composite (Fig. 11). Several systems have been explored with dendrimers, e.g., in the group of Crooks dendrimer-encapsulated metal nanoparticles were prepared in fluorous media, leading to easy separation and recycling of the catalyst. Chechik and Crooks describe a case where Pd nanoparticles in PAMAM dendrimers were complexed with carboxylic end groups of perfluoropolyethers in fluorous solvents.[80] Hydrogenation of various alkenes, including 1-hexene and methylacrylate, was performed to assess the catalytic activity of the system. Pd–dendrimer nanocomposites were observed to be catalytically active without appreciable loss of activity after 12 cycles. The wide range of turnover frequency numbers was observed exhibiting the possibility of selective hydrogenation; more polar substrates were hydrogenated faster because of the hydrophilic interior of the dendrimer that promotes contact between the metal nanoparticles and the polar substrates.

Yeung and Crooks also presented examples of carbon–carbon bond formation in biphasic fluorous/organic solvent systems catalyzed by dendrimer-encapsulated nanoparticles.[81] The perfluorinated polyether-derivatized poly(propylene imine) (PPI) dendrimers containing Pd(0) nanoparticles effectively catalyzed Heck couplings of aryl halides with *n*-butylacrylate. Because PPI dendrimers are more stable at elevated temperatures than PAMAM dendrimers, they are better suited for the Heck reaction, which requires high temperatures and long reaction times. Whereas the selectivity for the production of *n*-butyl-*trans*-formylcinnamate was previously reported to be in the 74–98% range in the absence of the dendritic stabilizer,[82] the PPI–Pd(0) catalyst provides 100% selectivity for the trans isomer. This result suggested that the sterically confined environment of the dendritic interior may impart regioselectivity.[83,84] It should be noted that the activity of catalysts can also be influenced by the dendrimer size. Equal amounts of metal nanoparticles of nearly the same particle size could be prepared within dendrimers of varying generations. A fifth-generation dendrimer/Pd(0) catalyst gave substantially greater

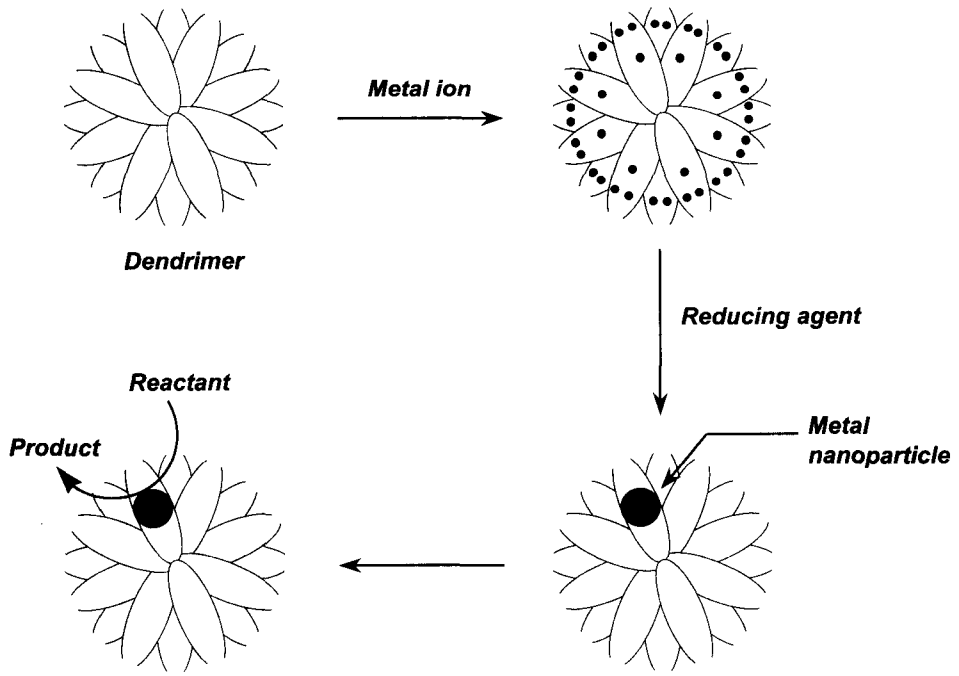


Fig. 11 Schematic illustration of the synthesis of metal nanoparticles within a dendrimer template.

product yield than the fourth-generation catalyst, which Crooks attributes to the larger dendrimer host that provides a less confining space for the Pd nanoparticles, resulting in easier access of substrate to the particle.[81]

Dendrimer–nanoparticle composites have been used in conjunction with supercritical CO_2, an alternative to hydrocarbon solvents, and a "greener" approach to catalysis with nanoparticles. Crooks has previously shown that Pd nanoparticles, encapsulated by PPI dendrimers having covalently grafted perfluorinated-polyether chains, were easily solubilized in this supercritical medium without requiring perfluorinated ligands or toxic solvents.[84] Whereas the standard Pd-catalyzed Heck coupling between iodobenzene and methylacrylate resulted in the *cis*- and/or *trans*-cinnamate product, the CO_2-soluble Pd-based nanocomposite exclusively produced methyl 2-phenylacrylate. The high selectivity is presumably related to both the steric environment of the dendrimer template and the properties of supercritical CO_2 because the same dendrimer-encapsulated Pd(0) yielded only the *trans*-cinnamate product using organic/fluorocarbon solvent system instead of supercritical CO_2.[81]

Amphiphilic block copolymers have been used for the stabilization of metal nanoparticles.[85–88] They are versatile components for polymer–metal systems because they can offer several options for tuning the system in catalytic applications. PS-*b*-PMMA and PS-*b*-PEO block copolymers were observed to be suitable for controlling both metal size and size distribution. The catalytic activities of the palladium and platinum nanoparticles stabilized by block copolymers were tested by hydrogenation of cyclohexene, which provided 100% conversion. The nanoparticle morphologies and catalytic activities were strongly dependent on many factors, including the type and ratio of polymer blocks, choice of metal precursor type, and the reduction method and conditions. These findings suggested that amphiphilic block copolymers could offer a tremendous versatility to the design of tailored catalysts.

Galow and coworkers have reported electrostatic-mediated assembly of highly reactive heterogeneous catalysts in which mixed monolayer protected clusters (MMPCs) were used as both building blocks and active catalysts.[89] Catalysts were formed through calcinations of the mixture of palladium colloid and silica–polymer aggregates (Fig. 12). A series of hydrogenation reactions and Heck reaction were performed and substantially high catalytic ability was observed. In the case of the Heck reaction between nitroarene and styrene, only 0.045 mol% of Pd was required to show a dramatic improvement over commercial counterparts.

The fabrication of nanoparticle-based templates that exhibit site-isolated nanoparticles in long-range ordered arrays may be ideal for some catalytic applications; thus future efforts in catalysis will be directed toward constructs as described in Fig. 9 above. Such studies,

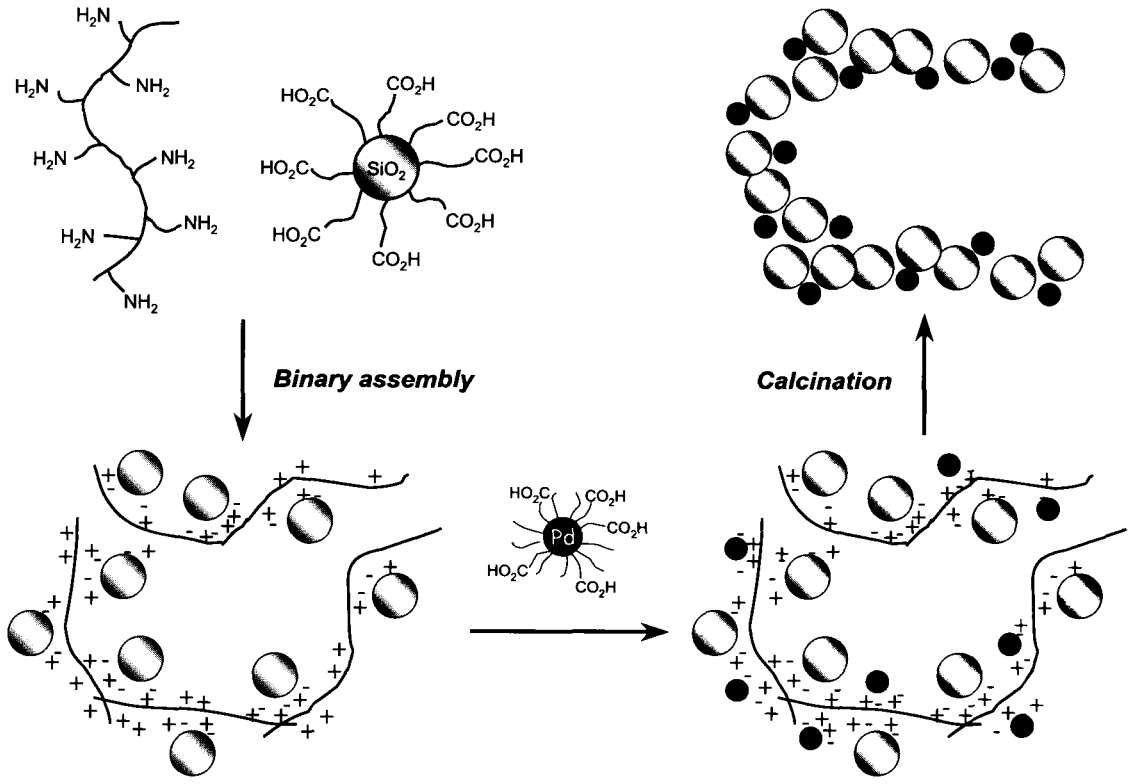


Fig. 12 Formation of palladium and silica nanoparticle catalysts directed by the interactions of functionalized polymers and nanoparticles. (*View this art in color at www.dekker.com.*)

performed previously with CdSe nanoparticles, should translate well to other nanoparticle systems studied in catalysis. Because selectivity is not understood as well as activity, future research should focus on achieving exceedingly high selectivity for the desired product. Through systematic variation of particle size and interparticle distance by using block copolymer templates with different molecular weight, structural ingredients responsible for selectivity of catalyst may be investigated in the near future.

CONCLUSION

In summary, it is clear that nanoparticle–polymer composite materials are playing an increasingly important role in nanotechnology. This article has summarized only a few of the many outstanding achievements in the field. As our ability to prepare and assemble these composites continues to improve, the specificity of the applications that can be addressed will increase accordingly. Nevertheless, progress to date has been most impressive, and all indications point toward further growth in the future.

ACKNOWLEDGMENTS

T.E. gratefully acknowledges the contributions of his group in nanoparticle–polymer composite materials, including Kevin Sill, Dr. Seunghoo Yoo, Habib Skaff, Rui Hong, Qingling Zhang, and Korisha Ramdhanie. Financial support for T.E.'s research group in nanoparticle–polymer composites was provided by the National Science Foundation (CAREER grant #0239486), the Eastman Kodak Company, and the U.S. Army Center of Excellence at U.Mass. Amherst.

REFERENCES

1. Goodyear, C. Improvement in India-Rubber Fabrics. US Patent 3633, June 15, 1844.
2. Goodyear, C. Procedure for the fabrication of objects which are more or less hard, bendable, or elastic from natural or synthetic rubber in conjunction with other materials. Dinglers Polytech. J. **1856**, *CXXXIX*, 376–390.
3. Goodyear, N. Improvement in the Manufacture of India-Rubber. US Patent 8075, May 6, 1851.

4. Baekeland, L.H. Bakelite, a new composition of matter: Its synthesis, constitution, and uses. Sci. Am., Suppl. **1909**, *68*, 322.
5. Baekeland, L.H. Method of Molding Articles. US Patent 939,966, January 29, 1909.
6. Sheldon, R.P. *Composite Polymeric Materials*; Applied Science Publishers Ltd.: Essex, England, 1982.
7. Shipway, A.N.; Katz, E.; Willner, I. Nanoparticle arrays on surfaces for electronic, optical, and sensor applications. Chem. Phys. Chem. **2000**, *1*, 18–52.
8. Gudiksen, M.S.; Lauhon, L.J.; Wang, J.; Smith, D.C.; Lieber, C.M. Growth of nanowire superlattice structures for nanoscale photonics and electronics. Nature **2002**, *415*, 617–620.
9. Service, R.F. Nanotechnology—Biology offers nanotechs a helping hand. Science **2002**, *298*, 2322–2323.
10. Williams, K.A.; Veenhuizen, P.T.M.; de la Torre, B.G.; Eritja, R.; Dekker, C. Nanotechnology—Carbon nanotubes with DNA recognition. Nature **2002**, *420*, 761.
11. Wang, F.C.; Zhang, W.E.; Yang, C.H.; Yang, M.J.; Bennett, B.R.; Wilson, R.A.; Stone, D.R. A tunneling field-effect transistor with 25 nm metallurgical channel length. Appl. Phys. Lett. **1997**, *70*, 3005–3007.
12. Conley, W.; Dean, K.; Miller, D.; Rich, G.; Graffenberg, V.; Patel, S.; Lin, S.H.; Jamieson, A.; Hung, R.; Yamada, S.; Pinnow, M.; MacDonald, S.; Chambers, C.; Osborne, B.; Patterson, K.; Somervell, M.; Trinque, B.; Tran, H.V.; Cho, S.; Chiba, T.; Byers, J.; Tomas, B.; Shafer, G.; DesMarteau, D.; Klopp, J.; Frechet, J.; Sanders, D.; Grubbs, R.; Ober, C.; Korner, H.; Willson, C.G. Developments in materials for 157 nm photoresists. Solid State Technol. **2001**, *44*, 63–68.
13. Black, C.T.; Guarini, K.W.; Milkove, K.R.; Baker, S.M.; Russell, T.P.; Tuominen, M.T. Integration of self-assembled diblock copolymers for semiconductor capacitor fabrication. Appl. Phys. Lett. **2001**, *79*, 409–411.
14. Mattoussi, H.; Mauro, J.M.; Goldman, E.R.; Green, T.M.; Anderson, G.P.; Sundar, V.C.; Bawendi, M.G. Bioconjugation of highly luminescent colloidal CdSe–ZnS quantum dots with an engineered two-domain recombinant protein. Phys. Status Solidi, B Basic Res. **2001**, *224*, 277–283.
15. Gerion, D.; Parak, W.J.; Williams, S.C.; Zanchet, D.; Micheel, C.M.; Alivisatos, A.P. Sorting fluorescent nanocrystals with DNA. J. Am. Chem. Soc. **2002**, *124*, 7070–7074.
16. Jaiswal, J.K.; Mattoussi, H.; Mauro, J.M.; Simon, S.M. Long-term multiple color imaging of live cells using quantum dot bioconjugates. Nat. Biotechnol. **2003**, *21*, 47–51.
17. Zerda, A.S.; Caskey, T.C.; Lesser, A.J. Highly concentrated, intercalated silicate nanocomposites: Synthesis and characterization. Macromolecules **2003**, *36*, 1603–1608.
18. Kojima, Y.; Usuki, A.; Kawasumi, M.; Okada, A.; Fukushima, Y.; Kurauchi, T.; Kamigaito, O. Mechanical-properties of nylon 6–clay hybrid. J. Mater. Res. **1993**, *8*, 1185–1189.
19. Kojima, Y.; Usuki, A.; Kawasumi, M.; Okada, A.; Kurauchi, T.; Kamigaito, O. Synthesis of nylon-6–clay hybrid by montmorillonite intercalated with epsilon-caprolactam. J. Polym. Sci., Part A, Polym. Chem. **1993**, *31*, 983–986.
20. Usuki, A.; Kojima, Y.; Kawasumi, M.; Okada, A.; Fukushima, Y.; Kurauchi, T.; Kamigaito, O. Synthesis of nylon 6–clay hybrid. J. Mater. Res. **1993**, *8*, 1179–1184.
21. Vaia, R.A.; Ishii, H.; Giannelis, E.P. Synthesis and properties of 2-dimensional nanostructures by direct intercalation of polymer melts in layered silicates. Chem. Mater. **1993**, *5*, 1694–1696.
22. Manias, E.; Touny, A.; Wu, L.; Strawhecker, K.; Lu, B.; Chung, T.C. Polypropylene/montmorillonite nanocomposites. Review of the synthetic routes and materials properties. Chem. Mater. **2001**, *13*, 3516–3523.
23. Porter, D.; Metcalfe, E.; Thomas, M.J.K. Nanocomposite fire retardants—A review. Fire Mater. **2000**, *24*, 45–52.
24. Carrodo, K.A. Synthetic organo- and polymer–clays: Preparation, characterization, and materials applications. Appl. Clay Sci. **2000**, *17*, 1–23.
25. Caskey, T.; Lesser, A.J.; McCarthy, T.J. Supercritical CO_2 welding of laminated linear low density polyethylene films. Polym. Eng. Sci. **2001**, *41*, 2259–2265.
26. Kung, E.; Lesser, A.J.; McCarthy, T.J. Morphology and mechanical performance of polystyrene/polyethylene composites prepared in supercritical carbon dioxide. Macromolecules **1998**, *31*, 4160–4169.
27. Balazs, A.C.; Singh, C.; Zhulina, E.; Lyatskaya, Y. Modeling the phase behavior of polymer/clay nanocomposites. Acc. Chem. Res. **1999**, *32*, 651–657.
28. Varghese, H.; Bhagawan, S.S.; Thomas, S. Effects of blend ratio, crosslinking systems and fillers on the morphology, curing behavior, mechanical properties, and failure mode of acrylonitrile butadiene rubber and poly(ethylene-*co*-vinyl acetate) blends. J. Appl. Polym. Sci. **1999**, *71*, 2335–2364.
29. Alivisatos, A.P. Perspectives on the physical

chemistry of semiconductor nanocrystals. J. Phys. Chem. **1996**, *100*, 13226–13239.
30. Peng, Z.A.; Peng, X. Formation of high-quality CdTe, CdSe, and CdS nanocrystals using CdO as precursor. J. Am. Chem. Soc. **2001**, *123*, 183–184.
31. Skaff, H.; Ilker, M.F.; Coughlin, E.B.; Emrick, T.S. Preparation of cadmium selenide–polyolefin composites from functional phosphine oxides and ruthenium-based metathesis. J. Am. Chem. Soc. **2002**, *124*, 5729–5733.
32. Thompson, R.B.; Ginzburg, V.V.; Matsen, M.W.; Balazs, A.C. Block copolymer-directed assembly of nanoparticles: Forming mesoscopically ordered hybrid materials. Macromolecules **2002**, *35*, 1060–1071.
33. Murray, C.B.; Kagan, C.R.; Bawendi, M.G. Self-organization of CdSe nanocrystallites into three-dimensional quantum dot superlattices. Science **1995**, *270*, 1335–1338.
34. Sooklal, K.; Hanus, L.H.; Ploehn, H.J.; Murphy, C.J. A blue-emitting CdS/dendrimer nanocomposite. Adv. Mater. **1998**, *10*, 1083–1087.
35. Brus, L. Chemical approaches to semiconductor nanocrystals. J. Phys. Chem. Solids **1998**, *59*, 459–465.
36. Lee, J.; Sundar, V.C.; Heine, J.R.; Bawendi, M.G.; Jensen, K.F. Full color emission from II–VI semiconductor quantum dot–polymer composites. Adv. Mater. **2000**, *12*, 1102–1105.
37. Corbierre, M.K.; Cameron, N.S.; Sutton, M.; Mochrie, S.G.J.; Lurio, L.B.; Ruhm, A.; Lennox, R.B. Polymer-stabilized gold nanoparticles and their incorporation into polymer matrices. J. Am. Chem. Soc. **2001**, *123*, 10411–10412.
38. Skaff, H.; Emrick, T. The use of 4-substituted pyridines to afford amphiphilic, pegylated cadmium selenide nanoparticles. Chem. Commun. **2003**, 52–53.
39. Carrot, G.; Scholz, S.M.; Plummer, C.J.G.; Hilborn, J.G.; Hedrick, J.L. Synthesis and characterization of nanoscopic entities based on poly(caprolactone)-grafted cadmium sulfide nanoparticles. Chem. Mater. **1999**, *11*, 3571–3577.
40. Lowe, A.B.; Sumerlin, B.S.; Donovan, M.S.; McCormick, C.L. Facile preparation of transition metal nanoparticles stabilized by well-defined (Co)polymers synthesized via aqueous reversible addition–fragmentation chain transfer polymerization. J. Am. Chem. Soc. **2002**, *124*, 11562–11563.
41. Watson, K.J.; Zhu, J.; Nguyen, S.T.; Mirkin, C.A. Hybrid nanoparticles with block copolymer shell structures. J. Am. Chem. Soc. **1999**, *121*, 462–463.
42. Skaff, H.; Ilker, M.F.; Coughlin, E.B.; Emrick, T. Preparation of cadmium selenide–polyolefin composites from functional phosphine oxides and ruthenium-based metathesis. J. Am. Chem. Soc. **2002**, *124*, 5729–5733.
43. Jordan, R.; West, N.; Ulman, A.; Chou, Y.M.; Nuyken, O. Nanocomposites by surface-initiated living cationic polymerization of 2-oxazolines on functionalized gold nanoparticles. Macromolecules **2001**, *34*, 1606–1611.
44. Zhou, Q.Y.; Wang, S.X.; Fan, X.W.; Advincula, R.; Mays, J. Living anionic surface-initiated polymerization (LASIP) of a polymer on silica nanoparticles. Langmuir **2002**, *18*, 3324–3331.
45. von Werne, T.; Patten, T.E. Preparation of structurally well-defined polymer–nanoparticle hybrids with controlled/living radical polymerizations. J. Am. Chem. Soc. **1999**, *121*, 7409–7410.
46. von Werne, T.; Patten, T.E. Atom transfer radical polymerization from nanoparticles: A tool for the preparation of well-defined hybrid nanostructures and for understanding the chemistry of controlled/"living" radical polymerizations from surfaces. J. Am. Chem. Soc. **2001**, *123*, 7497–7505.
47. Perruchot, C.; Khan, M.A.; Kamitsi, A.; Armes, S.P.; von Werne, T.; Patten, T.E. Synthesis of well-defined, polymer-grafted silica particles by aqueous ATRP. Langmuir **2001**, *17*, 4479–4481.
48. Farmer, S.C.; Patten, T.E. Photoluminescent polymer/quantum dot composite nanoparticles. Chem. Mater. **2001**, *13*, 3920–3926.
49. Nuss, S.; Bottcher, H.; Wurm, H.; Hallensleben, M.L. Gold nanoparticles with covalently attached polymer chains. Angew. Chem., Int. Ed. **2001**, *40*, 4016–4018.
50. Bottcher, H.; Hallensleben, M.L.; Nuss, S.; Wurm, H. ATRP grafting from silica surface to create first and second generation of grafts. Polym. Bull. **2000**, *44*, 223–229.
51. Mori, H.; Seng, D.C.; Zhang, M.F.; Muller, A.H.E. Hybrid nanoparticles with hyperbranched polymer shells via self-condensing atom transfer radical polymerization from silica surfaces. Langmuir **2002**, *18*, 3682–3693.
52. Frechet, J.M.J.; Henmi, M.; Gitsov, I.; Aoshima, S.; Leduc, M.R.; Grubbs, R.B. Self-condensing vinyl polymerization—An approach to dendritic materials. Science **1995**, *269*, 1080–1083.
53. Wang, Y.A.; Li, J.J.; Chen, H.Y.; Peng, X.G. Stabilization of inorganic nanocrystals by organic dendrons. J. Am. Chem. Soc. **2002**, *124*, 2293–2298.
54. Wang, R.Y.; Yang, J.; Zheng, Z.P.; Carducci, M.D.; Jiao, J.; Seraphin, S. Dendron-controlled nucleation

and growth of gold nanoparticles. Angew. Chem., Int. Ed. **2001**, *40*, 549–552.
55. Sun, S.H.; Murray, C.B. Synthesis of monodisperse cobalt nanocrystals and their assembly into magnetic superlattices (invited). J. Appl. Phys. **1999**, *85*, 4325–4330.
56. Thompson, R.B.; Ginzburg, V.V.; Matsen, M.W.; Balazs, A.C. Predicting the mesophases of copolymer–nanoparticle composites. Science **2001**, *292*, 2469–2472.
57. Lee, J.Y.; Thompson, R.B.; Jasnow, D.; Balazs, A.C. Effect of nanoscopic particles on the mesophase structure of diblock copolymers. Macromolecules **2002**, *35*, 4855–4858.
58. Lee, J.Y.; Thompson, R.B.; Jasnow, D.; Balazs, A.C. Entropically driven formation of hierarchically ordered nanocomposites. Phys. Rev. Lett. **2002**, *89*.
59. Fogg, D.E.; Radzilowski, L.H.; Dabbousi, B.O.; Schrock, R.R.; Thomas, E.L.; Bawendi, M.G. Fabrication of quantum dot–polymer composites: Semiconductor nanoclusters in dual-function polymer matrices with electron-transporting and cluster-passivating properties. Macromolecules **1997**, *30*, 8433–8439.
60. Boal, A.K.; Galow, T.H.; Ilhan, F.; Rotello, V.M. Binary and ternary polymer-mediated "bricks and mortar" self-assembly of gold and silica nanoparticles. Adv. Funct. Mater. **2001**, *11*, 461–465.
61. Boal, A.K.; Gray, M.; Ilhan, F.; Clavier, G.M.; Kapitzky, L.; Rotello, V.M. Bricks and mortar self-assembly of nanoparticles. Tetrahedron **2002**, *58*, 765–770.
62. Mamedov, A.A.; Belov, A.; Giersig, M.; Mamedova, N.N.; Kotov, N.A. Nanorainbows: Graded semiconductor films from quantum dots. J. Am. Chem. Soc. **2001**, *123*, 7738–7739.
63. Lopes, W.A. Nonequilibrium self-assembly of metals on diblock copolymer templates. Phys. Rev., E **2002**, *65*, 031606.
64. Yuan, Y.J.; Hentze, H.P.; Arnold, W.M.; Marlow, B.K.; Antonietti, M. Fabrication of nanostructured silica using a triblock copolymer template. Nano Lett. **2002**, *2*, 1359–1361.
65. Misner, M.J.; Skaff, H.; Emrick, T.; Russell, T.P. Directed deposition of nanoparticles using diblock copolymer templates. Adv. Mater. **2003**, *15*, 221–224.
66. Mansky, P.; Russell, T.P.; Hawker, C.J.; Pitsikalis, M.; Mays, J. Ordered diblock copolymer films on random copolymer brushes. Macromolecules **1997**, *30*, 6810–6813.
67. Huang, E.; Pruzinsky, S.; Russell, T.P.; Mays, J.; Hawker, C.J. Neutrality conditions for block copolymer systems on random copolymer brush surfaces. Macromolecules **1999**, *32*, 5299–5303.
68. Sill, K.; Yoo, S.H. Unpublished results.
69. Firestone, M.A.; Williams, D.E.; Seifert, S.; Csencsits, R. Nanoparticle arrays formed by spatial compartmentalization in a complex fluids. Nano Lett. **2001**, *1*, 129–135.
70. Zaera, F.; Somorjai, G.A. Hydrogenation of ethylene over platinum (111) single-crystal surfaces. J. Am. Chem. Soc. **1984**, *106*, 2288–2293.
71. Yeates, R.C.; Turner, J.E.; Gellman, A.J.; Somorjai, G.A. The oscillatory behavior of the CO oxidation reaction at atmospheric pressure over platinum single crystals: Surface analysis and pressure dependent mechanisms. Surf. Sci. **1985**, *149*, 175–190.
72. Goodman, D.W.; Kelly, R.D.; Madey, T.E.; Yates, J.T. Kinetics of the hydrogenation of CO over a single crystal nickel catalyst. J. Catal. **1980**, *63*, 226–234.
73. Spencer, N.D.; Somorjai, G.A. Iron single crystals as ammonia synthesis catalysts: Effect of surface structure on catalyst activity. J. Catal. **1982**, *74*, 129–135.
74. Dwyer, D.J.; Hardenbergh, J.H. The catalytic reduction of carbon monoxide over iron surfaces: A surface science investigation. J. Catal. **1984**, *87*, 66–76.
75. Youchang, X.; Naasz, B.M.; Somorjai, G.A. Alcohol synthesis from CO and H_2 over molybdenum sulfide. The effect of pressure and promotion by potassium carbon. Appl. Catal. **1986**, *27*, 233–241.
76. Gellman, A.J.; Farias, M.H.; Somorjai, G.A. The catalytic hydrodesulfurization of thiophene on the Mo (100) crystal surface. J. Catal. **1984**, *88*, 546–548.
77. Somorjai, G.A. The evolution of surface chemistry. A personal view of building the future on past and present accomplishments. J. Phys. Chem., B **2002**, *106*, 9201–9213.
78. Li, Y.; Hong, X.M.; Collard, D.M.; El-Sayed, M.A. Suzuki cross-coupling reactions catalyzed by palladium nanoparticles in aqueous solution. Org. Lett. **2000**, *2*, 2385–2388.
79. Li, Y.; El-Sayed, M.A. The effect of stabilizers on the catalytic activity and stability of Pd colloidal nanoparticles in the Suzuki reactions in aqueous solution. J. Phys. Chem. B **2001**, *105*, 8938–8943.
80. Chechik, V.; Crooks, R.M. Dendrimer-encapsulated Pd nanoparticles as fluorous phase-soluble catalysts. J. Am. Chem. Soc. **2000**, *122*, 1243–1244.
81. Yeung, L.K.; Crooks, R.M. Heck heterocoupling within a dendritic nanoreactor. Nano Lett. **2001**, *1*, 14–17.
82. Moineau, J.; Pozzi, G.; Quici, S.; Sinou, D.

Palladium-catalyzed Heck reaction in perfluorinated solvents. Tetrahedron Lett. **1999**, *40*, 7683–7686.

83. Chow, H.F.; Mak, C.C. Dendritic bis(oxazoline)-copper(II) catalysts. 2. Synthesis, reactivity, and substrate selectivity. J. Org. Chem. **1997**, *62*, 5116–5127.
84. Yeung, L.K.; Lee, C.T.; Johnston, K.P.; Crooks, R.M. Catalysis in supercritical CO_2 using dendrimer-encapsulated palladium nanoparticles. Chem. Commun. **2001**, 2290–2291.
85. Mayer, A.B.R.; Mark, J.E. Transition metal nanoparticles protected by amphiphilic block copolymers as tailored catalyst systems. Colloid Polym. Sci. **1997**, *275*, 333–340.
86. Mayer, A.B.R.; Mark, J.E.; Morris, R.E. Palladium and platinum nanocatalysts protected by amphiphilic block copolymers. Polym. J. **1998**, *30*, 197–205.
87. Mayer, A.B.R.; Mark, J.E. Immobilization of palladium nanoparticles on latex supports and their potential for catalytic applications. Angew. Makromol. Chem. **1999**, *268*, 52–58.
88. Mayer, A.B.R.; Mark, J.E. Poly(2-hydroxyalkyl methacrylates) as stabilizers for colloidal noble metal nanoparticles. Polymer **2000**, *41*, 1627–1631.
89. Galow, T.H.; Drechsler, U.; Hanson, J.A.; Rotello, V.M. Highly reactive heterogeneous heck and hydrogenation catalysts constructed through 'bottom-up' nanoparticle self-assembly. Chem. Commun. **2002**, 1076–1077.

Polypropylene and Thermoplastic Olefin Nanocomposites

Francis M. Mirabella, Jr.
Equistar Chemicals, LP, Cincinnati, Ohio, U.S.A.

INTRODUCTION

The purpose of this chapter is to review the progress toward the preparation of polypropylene/clay and thermoplastic olefin/clay nanocomposites. Further, the structure and properties of these type of nanocomposites are discussed and, where possible, structure/property relationships are described. Lastly, the progress toward the achievement of commercially viable nanocomposites of this type is assessed.

Polyolefins account for about 185 billion pounds per year or about 50% of the total worldwide plastic's industry production. Polypropylene (PP) accounts for about 20% of this total or 74 billion pounds per year production. Therefore polypropylene is one of the most widely used thermoplastics not only because of its balance of physical and mechanical properties, but also because of its environmental friendliness (e.g., recyclability) and low cost. The most common type is isotactic PP, while syndiotactic PP is a minor contributor to the PP market. Polypropylene has an attractive combination of low density, high stiffness and toughness, and heat distortion temperature above 100°C, which endows it with extraordinary versatility of properties and applications ranging from molded parts to films and fibers. Innovative catalyst and process technologies have significantly simplified its production, resulting in the minimization of catalyst residues, waxy byproducts, and low stereoregularity components, by the use of high activity and highly stereoselective catalysts, such as Ziegler–Natta and metallocene catalysts. Highly energy-efficient processes, such as gas phase and liquid-pool processes, have minimized the need for solvents. Polypropylene offers ready means for recycling by remolding or feedstock recovery or by recovery of its energy content in power plants and steel mills. It has also continued to make inroads into the engineering plastics markets by replacing more expensive and less environmentally friendly resins. The addition of mineral fillers to polypropylenes has widened its sphere of applications by enhancing its stiffness properties. However, the concomitant effect is typically a decrease in toughness properties because the inorganic filler particles act as stress concentrators.

Montmorillonite (mmt) is the most common clay mineral used in polymer/clay hybrids and is composed of an octahedral alumina sheet sandwiched between two tetrahedral silica sheets. The anionically charged sheets, or clay platelets, are strongly held together by cations, such as Na^+, Li^+, Ca^{2+}, Fe^{2+}, and Mg^{2+}. The tightly bound stacks of clay platelets may be readily separated by simple dissolution in a polar solvent, such as water. In addition, the natural clay is miscible with polar polymers in which the platelets readily disperse. However, separation of the clay platelets in a nonpolar polymer is difficult. The thickness of the individual clay platelets is about 1 nm and the gallery spacing (spacing is defined in terms of d-spacing from X-ray diffraction and small-angle X-ray scattering, which is the distance from one clay platelet center to the next) of the platelets in natural clay is about 1 nm. Organic modification of the clay with alkyl ammonium chlorides, in which the alkyl chain is typically composed of 6 to 15 carbons, expands the gallery spacing to about 2.2 to 2.4 nm in typical commercial organoclays. Addition of maleic anhydride-modified PP often results in intercalated clay with gallery spacing of greater than 3 nm. The typically desired state is, upon addition of PP, the complete dispersion, i.e., exfoliation and dispersal, of the clay platelets so that there is no correlation between platelets and therefore no maxima in the X-ray scattering pattern. Considering that clay is composed of platelets that are about 1 nm in thickness and around 1000 nm in lateral dimensions and assuming a density of 2.65 g/cm^3 yields an upper limit of about 750 m^2/g surface area per gram for fully exfoliated clay. This is 5 to 10 times the surface area per gram achievable with ordinary fillers. Exfoliated clay in polymers was shown to endow unique properties to the system, such as increased Tg, increased stiffness, maintenance of impact strength, increased barrier to gases and liquids, increased flame retardancy, etc.

POLYPROPYLENE

Structure

The extraordinary property enhancements accessible through the formation of polymer/clay mineral hybrid systems were pioneered by Toyota workers about a decade ago. In a series of papers, the synthesis of nylon-6/clay

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013656

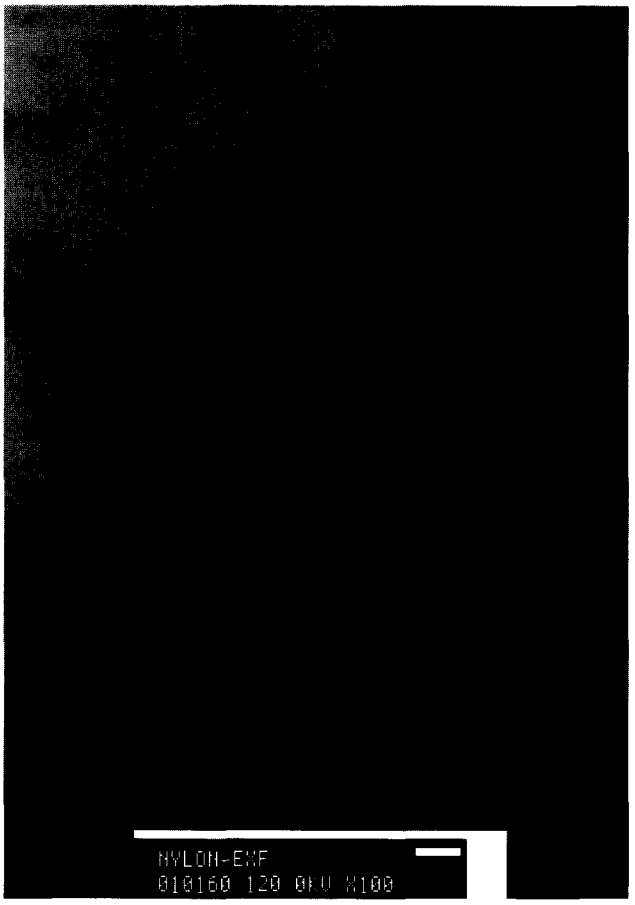


Fig. 1 Transmission electron microscopy photomicrograph of nylon/clay nanocomposite.

hybrids was described by the swelling of natural montmorillonite (fractionated from bentonite) with ε-caprolactam to form an intercalated compound, followed by the thermolytic polymerization of the ε-caprolactam to nylon-6.[1,2] The remarkable enhancements of the thermal and mechanical properties of the nylon-6 were further reported by the Toyota workers.[3] The dispersion of the silicate layers in such nylon-6/montmorillonite hybrids is generally excellent. The clay platelets are essentially completely exfoliated, resulting in a morphology as shown in the transmission electron microscopy (TEM) photomicrograph in Fig. 1. The clay platelets are seen to be completely dispersed and uncorrelated with each other. Therefore no X-ray signal from a periodic stacking of clay platelets would be expected from x-ray diffraction (XRD) or small angle x-ray scattering (SAXS) measurements.

In a technological advancement of this development, Vaia et al.[4] reported the formation of polymer/clay hybrids by the direct intercalation of polystyrene in organically modified montmorillonite. The montmorillonite was modified with alkyl ammonium cations to maximize the enthalpic interactions of the modified clay with the polystyrene and thus overcome the unfavorable loss of conformational entropy associated with the confinement of the polymer inside the clay galleries. The intercalated polymer was claimed to exist in a collapsed, two-dimensional conformation, and, as evidence of this, it was demonstrated that the normal glass transition temperature of the polystyrene was essentially absent for the intercalated polymer.[4] In further work, this group studied the chain dynamics of polystyrene in organically modified clay hybrids. They defined an effective diffusion coefficient and found that it increased with increasing alkyl chain length of the alkyl ammonium cation functional groups on the clay.[5]

The organic modification of natural clay minerals typically involves exchange of cations on the natural clay with alkyl ammonium cations. Commercial mmt typically is modified with hydrogenated tallow, which is ~65% C_{18}, 30% C_{16}, and 5% C_{14} alkyl chain lengths. The typical commercial organoclay is modified with an alkyl ammonium compound as shown in Fig. 2.

Two alternative synthetic routes were proposed by Manias et al.[6] to prepare organically modified mmt that would be readily dispersed in polypropylene. These small-scale studies were intended to explore the thermodynamics required to intercalate and exfoliate mmt with polypropylene. The first route relied on improving the favorability of the polypropylene–mmt enthalpic interactions to disperse the mmt. This was performed by functionalizing the polypropylene with 0.5–1.0 mol% of methyl styrene monomer. This random copolymer was then further functionalized with maleic anhydride and ethanol resulting in three random copolymers of polypropylene. The functionalized PP/dimethyl dioctadecyl-ammonium–mmt hybrid systems were formed by static melt intercalation in a vacuum oven at 180°C. The organo-mmt was observed by XRD to be intercalated with the functionalized PP and by TEM measurements to be dispersed in these random polypropylene copolymers to the extent of about 25–40% exfoliated clay platelets. The second route relied on decreasing the enthalpic interactions between the surfactant and the mmt. This is expected to effectively render the polymer–mmt interactions more

$(CH_3)_2$-N^+-HTCl$^-$

HT = hydrogenated tallow

HT = 65% C18, 30% C16 and 5% C14

Fig. 2 Structure of typical organic modifier for clay.

favorable. This was performed by the partial fluorination of the alkyl chains in an mmt. The mmt was organically modified with dimethyldioctadecyl ammonium and with partially fluorinated alkyltrichlorosilane. The organically modified mmt was statically melt intercalated with neat polypropylene. Measurements by XRD showed that the mmt was intercalated by the PP, indicating that the thermodynamics were favorable for nanocomposite formation. Further shear mixing of these two types of systems was expected to lead to further dispersion of the mmt in the polymer.

The promise of extraordinary improvement in end-use properties of polyolefin/clay hybrid systems has spurred great interest and intensive activity on the part of the players in the polyolefin and allied industries. Academic participation in these developments has also been vigorous. Optimistic predictions were made early in the drive to apply nanotechnology to the polyolefins industry. Probably, partly because of the promise of the technological benefits of nanotechnology and to the large size of the polyolefins industry, it was predicted that the largest impact of nanocomposites on the plastics industry would be in polyolefins. Wilson[7] of Inspired Innovations, LLC, predicted that nanocomposite versions of polyolefins will displace engineering thermoplastics and thermosets in many traditional applications and will open new opportunities to displace metal and glass in many traditional applications. Well over one-half of all the current applications for polymers could be improved by this new technology, according to Wilson. This author predicted millions of tons of nanocomposite polyolefin production and replacement of an equal amount of annual glass and metal production for the foreseeable future. Sinclair,[8] STA Research, predicted that polypropylene would be the largest growth polymer in nanocomposites and that PP nanocomposites would reach a volume of millions of tons per year by 2010. Both of these prognosticators emphasized that complete exfoliation of the clay was required to imbue the polyolefins with the improved end-use properties that could drive such massive commercial volumes of polyolefin nanocomposites. However, these goals could only be achieved by overcoming the very problematic dispersion of naturally polar clay in nonpolar polymers.

The attraction of clay mineral additives is their low cost to recover from the earth in their natural state. The incorporation of the clay in polypropylene, on the other hand, requires much more than natural clay. The technological challenges to formulating stable polypropylene/clay nanocomposites have proven to be formidable. Furthermore, the achievement of enhanced end-use properties has barely been explored because of this inability to form the systems with suitable clay dispersions.

Although studies continue aiming at novel modification of clay and functionalized polymer for improving the intercalation and exfoliation of the clay in polypropylene, the vast majority of experimental studies and commercial developments employ polypropylene with maleated polypropylene (PP-MA), or more rarely another functionalized PP, as a compatibilizer to form PP/clay nanocomposite hybrid systems. The approximate individual costs of organically modified mmt and PP-MA are about five times that of the polymer on a per pound basis. This means that the production of polypropylene nanocomposites with properties sufficiently enhanced to justify the additional costs is a doubly formidable problem.

In the earliest studies, the morphology of several systems was determined to compare various strategies to form intercalated and exfoliated polypropylene/clay hybrids systems. Low molecular weight polypropylene (12,000–30,000) functionalized with maleic anhydride or hydroxyl groups was found to intercalate into C_{18}-mmt.[9] Higher concentrations of functionalization were correlated with successful intercalation of the PP into the clay, while lower concentrations failed to exhibit intercalation of the PP. The same group then showed that PP homopolymer was not intercalated into Na-mmt, but a polyolefin diol was used to intercalate diastearyldimethylammonium-mmt followed by exfoliation of the clay by PP homopolymer.[10] In a further study, these workers employed PP-MA with two concentrations of maleic anhydride functionalization to intercalate C_{18}-mmt, followed by exfoliation of the clay with PP homopolymer.[11] It was found that too high a concentration of maleic anhydride functionalization resulted in immiscibility of the PP-MA compatibilizer with the PP homopolymer. The reinforcement of the system was judged best for the lower concentration of maleic anhydride in the PP-MA, which had a higher storage modulus and good miscibility with the PP homopolymer.

Many groups, to investigate strategies for improved clay dispersion in polypropylene and to probe the resulting morphologies, extended the foregoing early work by Okada et al. It was confirmed in some of these studies that PP homopolymer could not intercalate into organoclays.[12,13] The use of considerable amounts of PP-MA compatibilizer was necessary to intercalate the clay and to facilitate subsequent exfoliation of the clay platelets by PP homopolymer.[12,13] Xu et al.[12] found that C_{16}-mmt was intercalated in composites containing 2 wt.% clay only when the concentration of PP-MA (containing 0.9 wt.% maleic anhydride) exceeded 10 wt.% in the composite. Similarly, Walter et al.[13] found that C_{18}-synthetic clay was intercalated in polypropylene composites containing 7 wt.% clay and 20 wt.% PP-MA (containing 4.2 wt.% maleic anhydride), but that no intercalation of the PP occurred in the absence of the PP-MA. Fig. 3 shows TEM photomicrographs of a 7 wt.% clay composite with no PP-MA exhibiting no intercalation of the clay (Fig. 3A)

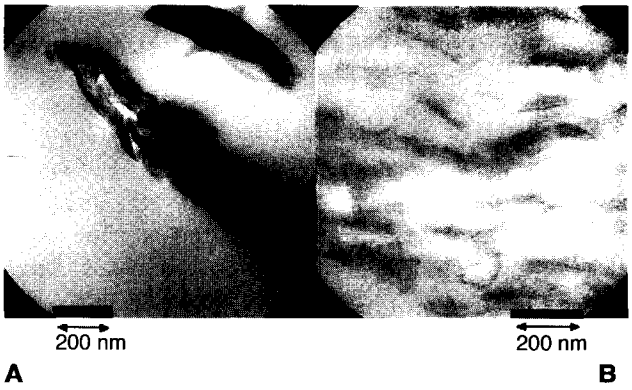


Fig. 3 Transmission electron microscopic image of polypropylene compound containing 10 vol.% organophilic fluoromica (M100/ODA) without addition of PP-g-MA compatibilizer (A, left) and the corresponding ME100/DA nanocomposite obtained in the presence of 20 wt.% PP-g-MA compatibilizer. (From Ref. [13] with permission from Marcel Dekker, Inc.)

and a 7 wt.% clay composite with 20 wt.% PP-MA exhibiting intercalation of the clay and some evidence of exfoliation of the clay "tactoids" by the PP homopolymer (Fig. 3B). Tactoids are defined as stacks of clay platelets that are intercalated with polymer chains, thereby expanding the galleries between succeeding clay platelets, as in Fig. 3B. Tactoids are distinguished from clay particles, which are composed of closely spaced clay platelets, which do not exhibit expanded galleries, as in Fig. 3A.

Reichert et al.[14] showed that the alkyl chain in clays organically modified with alkylammonium cations must exceed 8 carbon atoms in length to promote intercalation and that increasing the concentration of the maleic anhydride functionality in the PP-MA promoted exfoliated nanocomposite systems with PP homopolymer matrices. Alkyl chain lengths of C_4 to C_8 gave a similar clay interlayer distance of 1.3 nm, while C_{16} and C_{18} chain lengths afforded spacing of 1.7 and 2 nm, respectively (the nonmodified synthetic clay spacing was 0.95 nm). This indicated that sufficient molar mass of the alkyl chain in the organically modified clay was necessary to promote favorable interactions of the silicate organic modifier with the polymer chains intercalating the clay platelets. This type of observation was made in original work on organic modification of clay to promote intercalation of clay by nonpolar species.[15,16] Reichert et al.[17] further investigated the stability of polypropylene composites prepared with PP with Mw=350,000, 20 wt.% of PP-MA with Mw=32,000, and 4.2 wt.% MA and 5 wt.% of C_{18}-synthetic clay melt blended in a twin screw at 200°C. The as-prepared system exhibited typical nanocomposite morphology with well-intercalated tactoids and partial exfoliation of clay platelets in TEM photomicrographs, as well as a broad "hump" in the wide angle x-ray scattering (WAXS) pattern indicative of a broad distribution of interlayer distances. Upon annealing of the material at 220°C for 200 min, the morphology observed in the TEM appeared to show a collapse of the structure to much more tightly spaced stacks and no evidence of partial exfoliation of clay platelets. However, it was argued on the basis of WAXS data on this system that because no peak was observed and the broad "hump" observed for the as-prepared system disappeared in the annealed system, the clay exfoliation was further improved by the annealing process. The upper temperature limit of stability of the organic functionality, i.e., alkylammonium cations (C_{18}-ammonium cation in this case), on typical clays is about 190°C, as shown in Fig. 4.[18] Above this temperature, the organic functionality degrades and the intercalated clay tactoids characteristically collapse to a stacked platelet structure. This appears to be the case in this study and the TEM photomicrograph showing clay platelet stacks was strong evidence for collapse of the intercalated clay tactoids at the annealing temperature of 220°C.[17]

Many other studies of clay dispersion in polypropylene are available. Analytical methods to probe the morphology in such systems were discussed by Morgan and Gilman.[19] They recommended TEM as the best means to qualitatively assess the degree of dispersion of clay in polymer matrices. Their TEM photomicrographs demonstrated that unintercalated clay platelet stacks, intercalated tactoids, and exfoliated platelets could be recognized. They recommended XRD to measure d-spacing in ordered, unintercalated stacks and ordered, intercalated tactoids, but cautioned that XRD is insufficient for determining disordered and exfoliated clay because these exhibit no reflection peak. These analytical methods were used to probe the morphology in studies of melt intercalated PP/clay hybrid systems.[20–27] These studies all

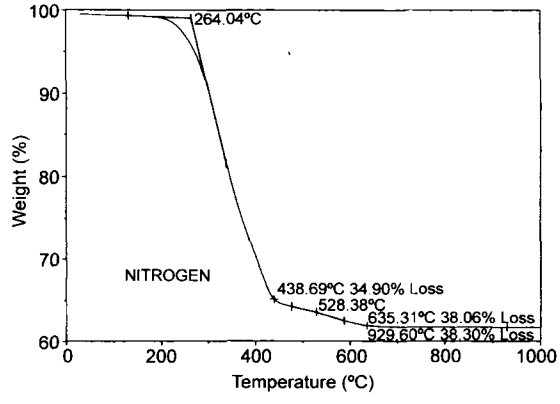


Fig. 4 Thermogravimetric analysis curve of Cloisite 20 A™ (Southern Clay Products) organo-mmt. The onset of thermal decomposition is at 190°C.

relied on significant concentrations of a functionalized polypropylene (predominantly PP-MA) to intercalate organically modified clay (predominantly alkylammonium-mmt) in a PP homopolymer matrix, followed by shearing of the melt to exfoliate the clay to the extent possible. In all cases, the best dispersion of the clay observed was partially as ordered, intercalated tactoids and partially as disordered, exfoliated clay platelets; however, nonuniform dispersion of the clay was commonly observed in all of these studies. The factors investigated were the MA concentration of the functionalized PP-MA and the concentration of PP-MA in the blend,[12,20,24,26,27] various melt compounding techniques,[12,20–22,25] and organic modification of clay.[23,24]

More uniform dispersion of alkylammonium-mmt in polypropylene was achieved by the intercalative polymerization of propylene in the galleries of the clay.[28] The TEM photomicrographs showed that the clay platelets were uniformly dispersed in the polypropylene. This strategy is attractive because no functionalized polypropylene compatibilizer is required; however, in situ polymerization of monomers in the clay galleries has not been demonstrated for commercial practice. Therefore melt intercalation remains the desired strategy for commercial production of polypropylene/clay hybrid nanocomposites. This goal has not been achieved for widespread commercial practice.

Mechanism of exfoliation

A mean-field, lattice-based model of polymer intercalation in organically modified silicates was presented by Vaia and Giannelis.[29,30] The thermodynamic factors which control polymer intercalation into the clay galleries were outlined in terms of an interplay between entropic and energetic factors. The entropic penalty for confining the polymer inside the clay gallery may be partially compensated for by an increase in entropy of the surfactant chains (alkyl chains on organic modifier) as the clay layers separate. Although the tethered alkyl chains cannot increase their entropy by an increase in translational freedom, the expanded gallery spacing permits increased conformational entropy. Intercalation will be thermodynamically possible when favorable polymer–organically modified silicate energetic interactions overcome any unfavorable entropic factors. Further separation, i.e., driving exfoliation, depends on maximization of favorable polymer–surface interactions with the organically modified silicate, as well as minimization of unfavorable interactions with the tethered alkyl chains of the organic modifier. This explains the use of polymer functionalization with polar moieties. The polar functionalities on polypropylene participate in polar–polar interactions with the silicate surface, decreasing the free energy of the system. If the total free energy of the system is decreased by the combined entropic and energetic factors, then intercalation and exfoliation may occur.

The further effects of shear forces on intercalated clay stacks (tactoids) have been elucidated by subsequent studies.[31,32] Fornes et al.[31] proposed a model explaining exfoliation of intercalated silicate stacks by the stepwise skewing of the silicate sheets in the tactoids, followed by peeling, one-by-one, of the silicate layers off the silicate stacks by combined diffusion/shearing, as shown in Fig. 5. The tactoids were shown to first skew apart forming shorter stacks of fewer and fewer numbers of silicate layers, followed by peeling of the individual layers off the tactoids forming exfoliated and dispersed layers, without short-range order with other silicate layers. The matrix polymer needs to have sufficient affinity for the silicate surface to cause spontaneous wetting of the surface to drive exfoliation. This is the required thermodynamic driving force. However, shear forces play a kinetic role by driving the peeling of silicate layers off the tactoids during

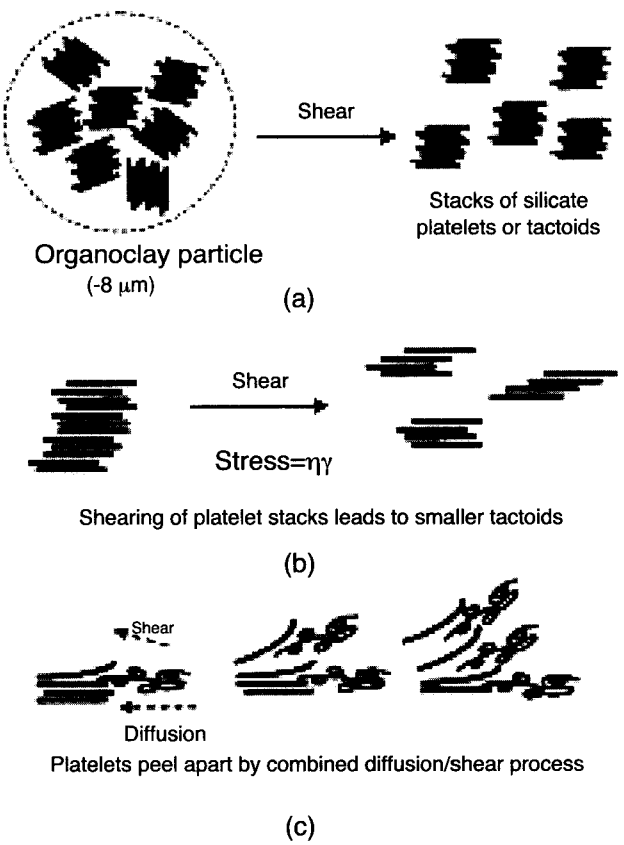


Fig. 5 Stepwise mechanism of clay platelet exfoliation in the melt compounding of nanocomposites: (a) organoclay particle breakup, (b) clay tactoid breakup, and (c) platelet exfoliation. (From Ref. [31] with permission from Elsevier Science Ltd.)

the residence time of the system in the shear process. It was found that the higher the melt viscosity of the matrix polymer, the greater the stress exerted on the tactoids, driving this skewing and peeling process. Therefore higher molecular weight matrix polymer was associated with higher stresses and a higher degree of exfoliation. This model was supported by a further experimental study in which the clay tactoid was observed to be in the process of peeling apart in TEM images, as shown in Fig. 6.[32] The silicate layers in the tactoid can be seen to be sliding and peeling apart from one another, like a deck of cards, in Fig. 6.

Crystallization

The effects of crystallization on polypropylene/clay hybrids have been investigated from two vantage points. The effects of the clay on the crystallization of the polypropylene matrix and the effects of the crystallization of the polypropylene on the clay morphology have been reported. The rate of crystallization was found to be increased for polypropylene-mmt nanocomposites as compared with neat polypropylene, as measured by crystallization half-times determined from differential scanning calorimetric data.[33,34] The nucleating effect of mmt has also been discussed in detail.[35,36] The nucleating effect was found to strongly depend on the origin and treatment of the mmt.[35] The spherulite diameter was decreased in PP-MA/mmt nanocomposites, as compared with neat polypropylene, in quiescent isothermal crystallization.[36]

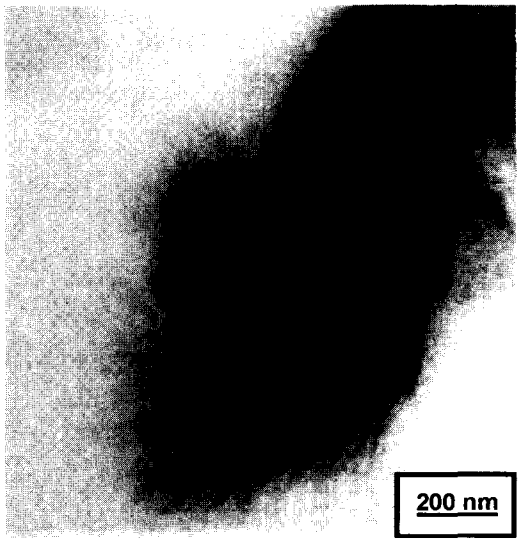


Fig. 6 Transmission electron micrograph of silicate layers peeling apart via a fan-shaped intermediate. (From Ref. [32] with permission from John Wiley and Sons, Inc.)

The quiescent, isothermal crystallization kinetics of PP/PP-MA/mmt nanocomposites were found to be slightly retarded in comparison to the neat polypropylene, but were not significantly different from the PP/PP-MA blend.[26] However, marked flow-induced acceleration of crystallization for PP/PP-MA/mmt nanocomposites was observed at strain rates for which flow had only a modest effect on the neat PP.[26] This behavior of the PP/PP-MA/mmt nanocomposites was further inferred to be largely due to the PP-MA compatibilizer, by comparing a blend of the PP/PP-MA, under these conditions.

The detailed crystalline morphology of the polypropylene matrix in PP/PP-MA/mmt nanocomposites has been reported.[36–38] Modifications observed were a fibrous crystalline morphology, instead of the usual spherulitic morphology after quiescent, isothermal crystallization.[38] Okamoto et al.[36,37] reported an interfibrillar structure composed of a disordered arrangement of cross-hatched lamellae among the dispersed clay platelets, which arose after quiescent, isothermal crystallization of a PP-MA/mmt nanocomposite because of crystallization occurring between clay particles, thereby disordering the usual spherulitic structure. Polarized optical microscopy, XRD, SAXS, and TEM measurements were correlated to propose a morphology in which fibrils of about 3–5 polypropylene lamellae reside between dispersed clay platelets, forming a disordered spherulitic superstructure.

Okamoto et al.[36,39] proposed an intercalation mechanism of PP-MA into mmt in PP-MA/mmt nanocomposites involving the interplay of crystallization and residence time in the melt state. Their results indicated that slower rate of crystallization permits more time for the polymer to intercalate the mmt by diffusing into the clay galleries, while faster crystallization rate inhibits this diffusion. They found that clay gallery spacing, measured by XRD, became larger as crystallization temperature was increased near the melting temperature. Interestingly, because crystallization is also slow just above glass transition temperature (Tg), it was similarly claimed, based on XRD measurements, that clay gallery spacing became larger for crystallization just above the glass transition temperature. In addition, gallery spacing was increased as clay concentration in the nanocomposite was decreased.

Mechanical Properties

A primary goal for nanocomposites is to achieve enhanced mechanical properties. The addition of inorganic fillers to polymers is widely practiced in the plastics industry. The common enhancement to polymer mechanical properties, by the addition of talc, calcium carbonate, glass fibers, wollastonite, glass beads, mica, wood flour, etc., is the increase in stiffness (Young's modulus. tensile modulus,

and flexural modulus), while strength properties (tensile strength, elongation at break, and flexural strength) and toughness (Izod and Charpy impact strength) may decrease or be maintained.[40,41] The degradation in strength and toughness properties is a result of the concentration, degree of dispersion, and particle size of the inorganic filler.[40,41] Typically, the degradation in properties is correlated with the size of the particulate because the particle acts as a stress concentrator, thereby initiating cracks and degrading strength and toughness properties.[40,41] For example, the decrease in notched Charpy impact strength vs. talc particle diameter is shown in Fig. 7.

The early promise of nanocomposites was to provide a type of inorganic filler that would enhance stiffness and strength properties, while simultaneously maintaining or even improving toughness properties.[1] This development was expected to be related to three unique structural characteristics of nanofiller particles: 1) their approximately 1-nm size in one dimension; 2) their large aspect ratio; and 3) their ability to be preferentially oriented because of the large aspect ratio.

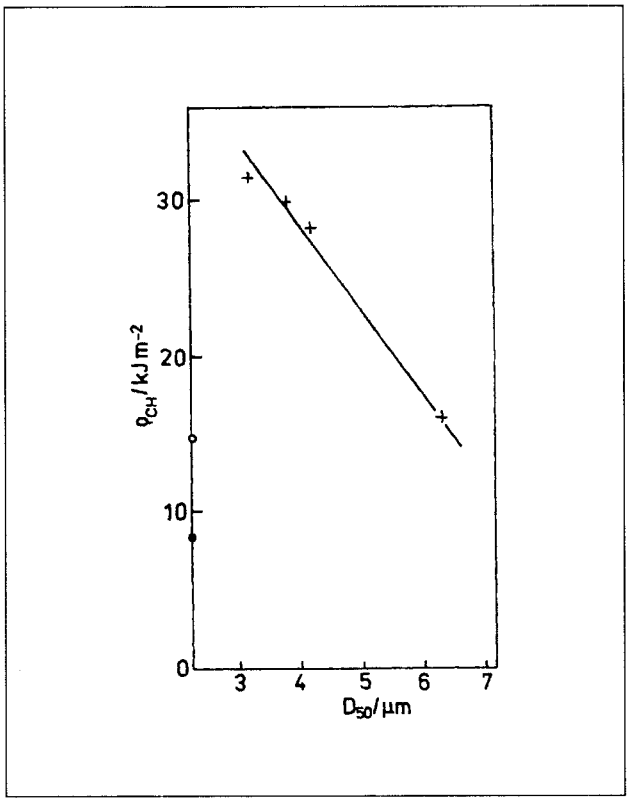


Fig. 7 Composite Charpy notch toughness vs. talc mean particle size of ternary composite (26 wt.% talc). (From Ref. [40] with permission from John Wiley and Sons, Inc.)

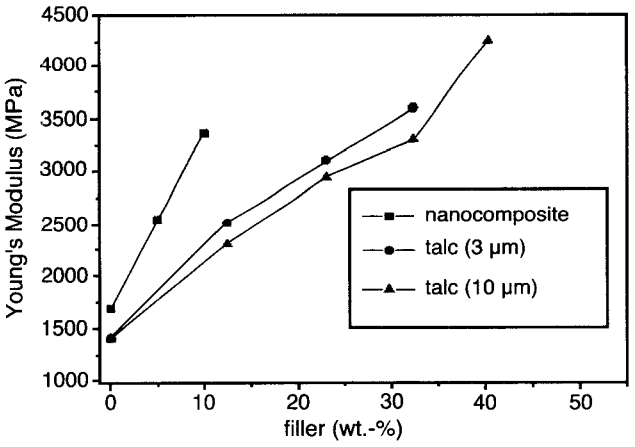


Fig. 8 Young's modulus of polypropylene compounds, prepared by melt compounding, as a function of filler content using organophilic fluoromica and talc with 3 and 10 µm average particle diameter. (From Ref. [13] with permission from Marcel Dekker, Inc.)

Another advantage, demonstrated by Walter et al.,[13] was the increase in matrix reinforcement achievable with nanocomposites with much lower inorganic filler concentration. The effectiveness of nanodispersed silicate relative to microdispersed talcs is shown in Fig. 8.

In the case of polypropylene, the further significant challenge is to achieve property enhancements that justify the incremental cost increase associated with the production of nanocomposites. This challenge is formidable because of the very low cost of polypropylene resins.

There have been relatively few studies on polypropylene nanocomposite properties and structure/property relationships in nanocomposites. This is partly because of the fact that the formation of anything approaching ideal PP nanocomposites has been unsuccessful. Therefore, as stated previously, most work has focused on the preparation and characterization of PP nanocomposites.

Early studies by Toyota workers, led by Hasegawa et al.,[42] focused on the increase of the dynamic storage modulus. PP/PP-MA/C18-mmt nanocomposites were formed, e.g., by blending 5 wt.% mmt and 22 wt.% PP-MA (MA acid number 52 mg KOH/g) with isotactic PP. It was found that only large concentrations of PP-MA afforded reasonable clay dispersion; however, even in these type blends partial exfoliation and partial intercalation (as intercalated tactoids of mmt) were observed in TEM measurements. In the nanocomposite blend cited above, the storage modulus was 1.8 times higher than that of the corresponding PP at 80°C. Hasegawa et al.[43] prepared PP-MA (0.2 wt.% MA)/C18-mmt (5.3 wt.% mmt)

nanocomposites, which exhibited apparently complete exfoliation of the mmt, and had 1.9 times higher tensile modulus than the PP-MA matrix polymer. However, a simple PP/C_{18}-mmt composite had a slightly higher tensile modulus than this PP-MA/C18-mmt nanocomposite.

Reichert et al.[14] made an extensive study of PP/PP-MA/synthetic clay compounds prepared with fluoromica-modified synthetic clays, organically modified with alkyl chains ranging in length from C_4 to C_{18}. These workers found that 10 wt.% synthetic clay modified with C_{16} organic modifier in conjunction with 20 wt.% PP-MA with 4.2 wt.% MA was sufficient to achieve effective PP matrix reinforcement. This nanocomposite exhibited exfoliated and well-dispersed silicate layers in the TEM and an increase in Young's modulus from 1490 to 3460 MPa and yield stress from 33 to 44 MPa with respect to bulk PP. However, notched Izod impact strength decreased from 1.7 to 1.4 kJ/m^2 with respect to bulk PP.

In other studies, polypropylene/clay nanocomposites exhibited enhancements in tensile strength and modulus[44] and also flexural strength and modulus.[45] Improvements in tensile properties were shown to be accompanied by maintenance or modest increases in fracture toughness.[46,47]

Monotonic increase in tensile modulus and decrease in elongation at break were observed for PP-MA/tallow-mmt nanocomposites up to 20 wt.% clay.[48] Tensile modulus increased rapidly up to about 5 wt.% clay, but then remained constant up to about 15 wt.% clay and then decreased at up to 20 wt.% clay. Notched impact strength remained approximately constant up to 20 wt.% clay, exhibiting modest increases between about 3 and 10 wt.%

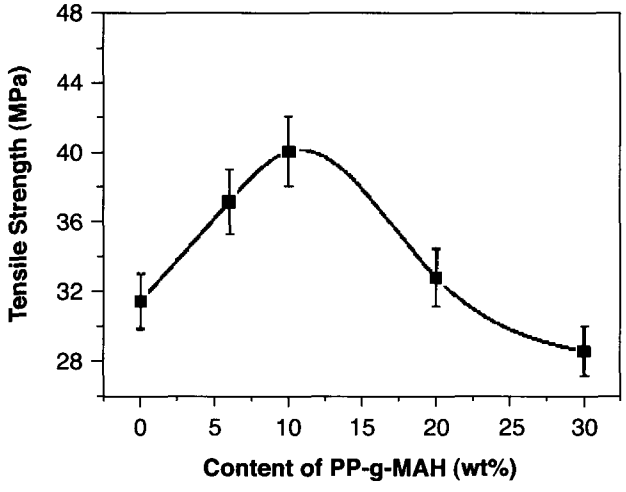


Fig. 9 Plot of tensile strength vs. concentration of PP-g-MAH. The mass percent of Org-MMT was 2 wt.%. (From Ref. [12] with permission from John Wiley and Sons, Inc.)

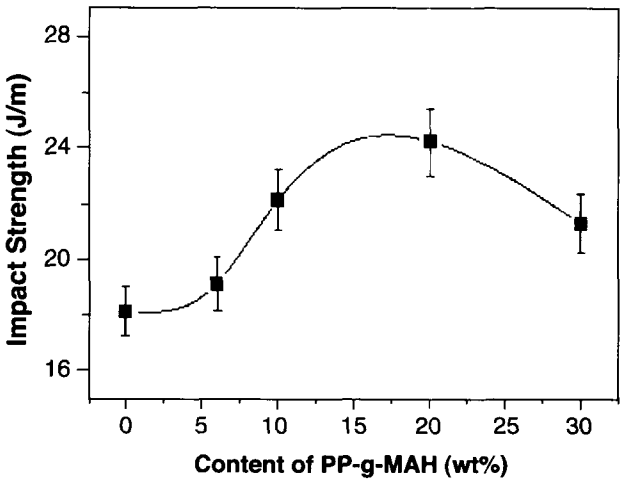


Fig. 10 Plot of impact strength vs. concentration of PP-g-MAH. The mass percent of Org-MMT was 2 wt.%. (From Ref. [12] with permission from John Wiley and Sons, Inc.)

clay. The best mechanical properties were observed for the highest molecular weight PP-MA.

In a study of PP/PP-MA/C_{16}-mmt nanocomposites, it was found that the best intercalation efficiency was achieved by the highest concentration of PP-MA (20–30 wt.%) and by a higher concentration of MA (0.9 wt.% better than 0.6 wt.%) in the PP-MA.[12] Tensile strength and notched Izod impact strength were found to increase and then decrease according to similar patterns, as shown in Figs. 9 and 10, as PP-MA concentration in the nanocomposite increased at 2 wt.% mmt concentration. Attempts to intercalate the same clay with the bulk polypropylene indicated little or no entry of the PP into the clay galleries.

The incorporation of radiation-grafted SiO_2 nanoparticles into polypropylene was shown to lead to well-dispersed SiO_2 nanoparticles grafted to the PP homopolymer matrix.[49] Reinforcement and toughening of the PP matrix were observed at concentrations between 0.5 and 5 wt.% SiO_2.

The micromechanics of nanocomposites reinforced with platelet-shaped fillers was addressed by the development of a theoretical model.[50] The model was aimed at predicting the buckling of platelets in reinforced materials under compressive loading. The model predicted a critical strain above which platelet buckling would lead to reduction in the compressive modulus relative to the tensile modulus. The model predicted a significant reduction in reinforcement efficiency as a result of incomplete exfoliation of platelets into "pseudoparticle" stacks containing polymer layers sandwiched between successive clay platelets (tactoids), rather than into individual perfectly

exfoliated and well-dispersed platelets. The model also predicted the reduction of reinforcement efficiency as a result of the deviation of the platelet from perfect biaxial in-plane orientation.

The common finding in these studies of the mechanical properties of polypropylene nanocomposites is that a large concentration of PP-MA is necessary to produce intercalated and exfoliated compounds with polypropylene. The nanocomposites exhibit imperfect dispersion of the clay with morphologies containing partial intercalated tactoids and partial exfoliated clay platelets. The improvements of polypropylene properties (tensile and toughness) are typically not sufficient to justify the additional costs to modify the polypropylene.

Rheology

The effects of clay and the orientation of the clay platelets on the rheology of polypropylene nanocomposites have been an area of particular interest, secondary to mechanical properties.[25,51–59] The rheological response of PP/PP-MA/mmt nanocomposites was studied.[58,59]

Lele et al. found that the zero-shear viscosity of the compatibilized nanocomposites (PP/PP-MA/tallow-mmt) was at least 3 orders of magnitude higher than that of the matrix PP and the uncompatibilized hybrids, as shown in Fig. 11. The strong dependence of melt viscosity on clay concentration was claimed to arise from frictional interactions between the silicate layers and not because of immobilization of confined chains between silicate layers. At low shear rates, a 3-D percolated structure of dispersed clay tactoids having a high zero-shear viscosity was postulated. At higher shear rates, this network was speculated to break, imparting a solid-like behavior to the melt, which exhibited yielding and decreasing viscosity as shear rate increased. Rheo-XRD experiments gave direct evidence of flow-induced orientation of the clay tactoids. Yielding was linked to the cooperative orientation of the clay. The rheological data for the uncompatibilized hybrid exhibited a much lower zero-shear viscosity and much less pronounced yield behavior, and the corresponding rheo-XRD data showed a smaller orientation of the clay tactoids.

This dramatic increase in low shear rate viscosity has been utilized to prepare polypropylene homopolymers and copolymers/clay compounds with improved melt strength in extrusion for the production of films, fibers, and a variety of molded articles.[60] The rheology of such clay compounds was shown to be controllable by the variation of the ratio of PP-MA compatibilizer/clay.

Barrier

The enhancement of barrier properties of polypropylene is obviously an important application for nanocomposites.

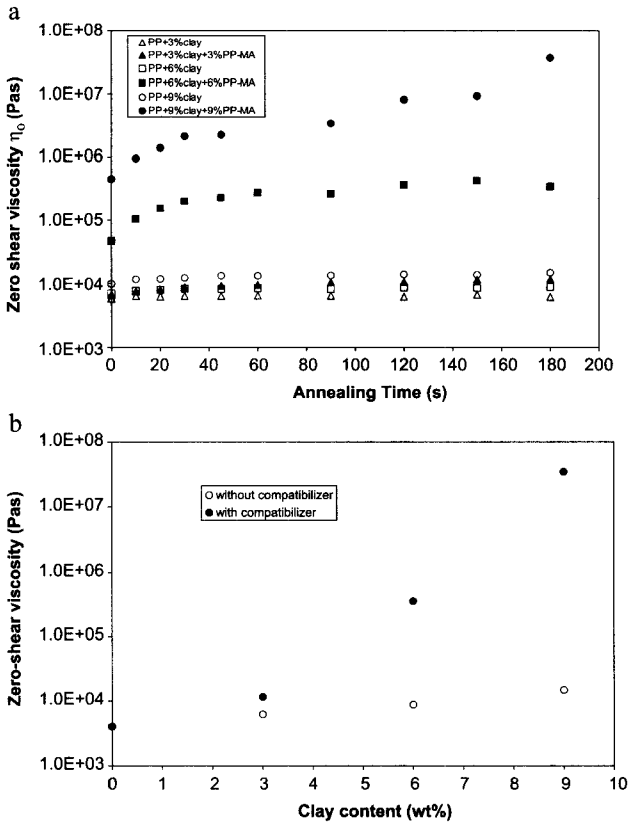


Fig. 11 (a) Zero shear viscosity as a function of annealing time for several PPCH samples in the presence (denoted by filled symbols) and absence (denoted by open symbols) of PP-MA. (b) Zero-shear viscosity as a function of clay content for samples with and without compatibilizer. (From Ref. [58] with permission from the American Chemical Society.)

The platelet morphology of dispersed clay would be expected to impart enhanced barrier properties to fabricated articles, and extraordinary barrier properties might be imparted by the preferential orientation of the platelets in such fabricated articles. Blow-molded containers, based on high-density polyethylene (HDPE)/mmt nanocomposites, were shown to have a much higher barrier to organic liquids compared with the neat HDPE.[61] The HDPE was melt-compounded with mmt, and the blow-molding process was optimized to obtain a high degree of orientation of the exfoliated clay platelets to maximize barrier performance. However, there have been only a few reports concerning the enhancement of barrier properties in polypropylene.[6,62] These workers reported that the permeability of PP/PP-MA/mmt or PP-MA/mmt nanocomposites, containing 4 wt.% mmt, was one-half of that of the matrix material, while solvent uptake was also decreased accordingly.

Flammability and Thermal Stability

Montmorillonite-based fillers have been shown to promote flame retardancy in polypropylene composites. The mechanism of this flame-retardant behavior has been shown to be based on the formation of a carbonaceous-char layer on the outer surface of the composite article by Gilman et al.[63] The surface char layer was found to have a high concentration of mmt layers and was claimed to become an excellent insulator and mass transport barrier, slowing oxygen diffusion into the surface and the escape of combustion products out of the surface. These workers showed that the heat release rate (HRR) was 75% lower for PP-MA/mmt (4 wt.%) nanocomposite compared with the neat PP-MA matrix (Fig. 12). In addition, the mass loss rate (MLR) of the same PP-MA/mmt (4 wt.%) nanocomposite was about one-quarter of that of the PP-MA matrix polymer (Fig. 13). These workers further found that the type of layered silicate, degree of dispersion, and the processing conditions affected the magnitude of the flammability reduction in the nanocomposites.

In another study, similar results were obtained and explained by a similar mechanism, based on the formation of a char layer.[64] A comparison was made between nanocomposites, based on a PP-MA matrix, containing 2.5 to 10 wt.% clay, with talc composites having the same PP-MA matrix. It was found that the limiting oxygen index (LOI, ASTM 2863-87) was increased for the PP-MA/clay nanocomposites over the talc composites (Fig. 14). Furthermore, the time from ignition to the first occurrence of dripping was shown to increase with clay loading (Fig. 15). Burning was reported to be much less intensive

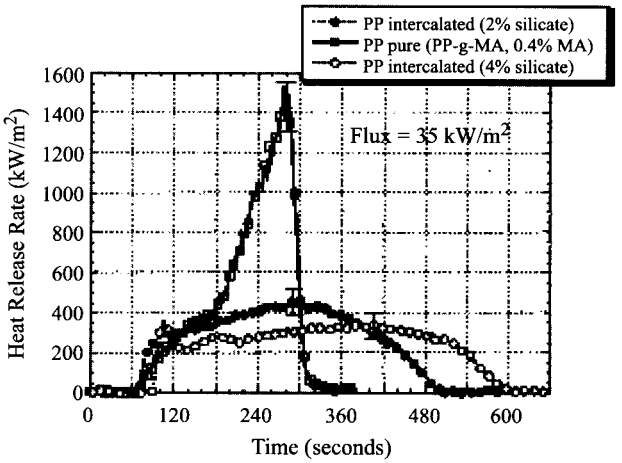


Fig. 12 Comparison of the heat release rate (HRR) for pure PPgMA and two PPgMA-layered silicate nanocomposites, at 35 kW/m^2 heat flux, showing a 70–80% reduction in peak HRR for the nanocomposites with a mass fraction of only 2 or 4 wt.% layered silicate, respectively. (From Ref. [63] with permission from the American Chemical Society.)

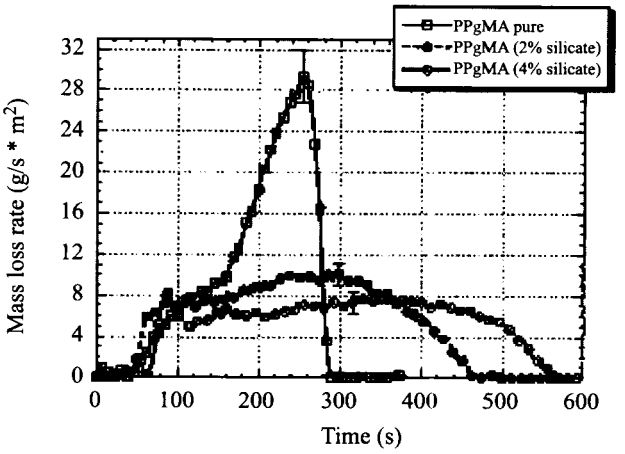


Fig. 13 Mass loss rate plots for PPgMA and two PPgMA-layered silicate nanocomposites. (From Ref. [63] with permission from the American Chemical Society.)

in the nanocomposites in comparison to the neat matrix polymer. In addition, ash layers of about 1 mm in thickness were formed on the nanocomposites, which were correlated with the minimization of burning intensity and dripping. In other studies, it was similarly reported that HRR and time-to-ignition were reduced in PP/clay nanocomposites.[6,65,66]

The detailed thermal behavior and the formation of char layers on PP/PP-MA/clay nanocomposites were reported.[67] The thermal oxidation process of the polymer was significantly slowed in the nanocomposites exhibiting high char yield. The inhibition of oxidation was explained

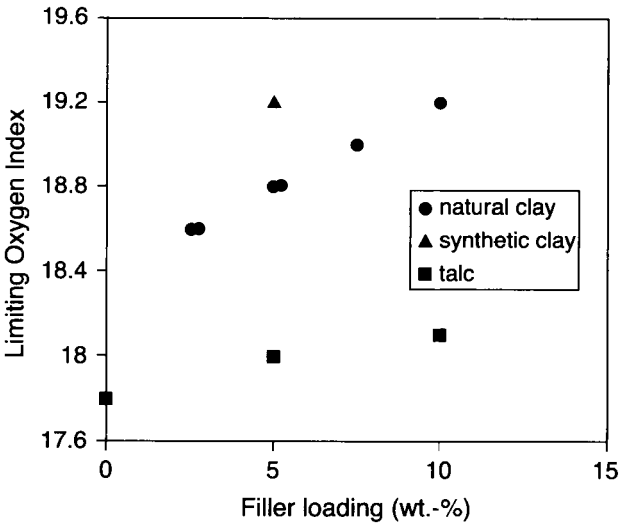


Fig. 14 Limiting oxygen index of composites with different inert fillers. (From Ref. [64] with permission from John Wiley and Sons, Inc.)

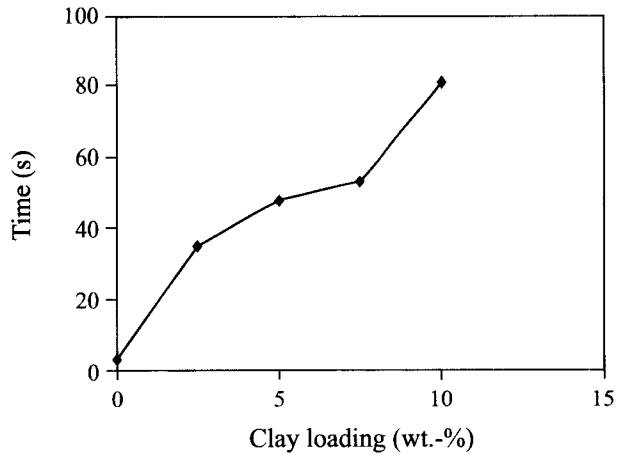


Fig. 15 Time from taking off the gas flame from the sample to the dripping of first burning parts. (From Ref. [64] with permission from John Wiley and Sons, Inc.)

on the basis of a physical barrier effect of the surface char layer, by ablative reassembling of the silicate and by a chemical catalytic action, because of the silicate and of the amine modifier on the silicate.

Heat Deflection Temperature

The heat deflection temperature (HDT) of a PP/fluoro-organo-mmt nanocomposite was found to increase from 109°C to 152°C compared to the neat PP.[6] The increase in HDT was attributed to better mechanical stability of the nanocomposite and not to an increase in melting temperature, which remained invariant in the nanocomposite relative to the neat PP.

Electrical Properties

A process to prepare electrically conducting PP-MA/graphite nanocomposites was reported.[68] It was found that the percolation threshold for conductivity was much lower than that of conventional conducting composites.

THERMOPLASTIC OLEFINS

Although the vast majority of work that has been published, relative to propylene-based polymers, concerns the development of polypropylene nanocomposites, there has been little incentive to commercially produce these compositions. Polypropylene is a very low cost commodity thermoplastic and the improvements in the properties so far demonstrated for polypropylene nanocomposites have not justified the increased cost to produce nanocomposites commercially.

On the other hand, far more interest has been generated in developing thermoplastic olefin (TPO) nanocomposites. The applications of TPOs, especially targeted for nanocomposite development, are those in the automotive sector. Thermoplastic olefins are usually reinforced with large quantities (up to 50 vol.%) of mineral fillers, such as talc and calcium carbonate. The outlook of replacing these conventional composites with lighter weight and better-performing nanocomposites has driven intensive efforts to achieve this objective. Because of the much higher cost basis of TPOs, relative to polypropylene homopolymer, this objective appears to be much easier to achieve.

Some relative disadvantages of PP are that it is brittle and has poor low-temperature properties. However, the addition of a rubber, such as ethylene-propylene rubber (EPR), or ethylene-propylene-diene-monomer (EPDM) results in dramatic improvements in the impact strength and low-temperature properties. These rubber-modified PPs have a greatly expanded sphere of applications, relative to PP, and are often called impact copolymer-polypropylene (ICP) and thermoplastic olefins (TPOs). Rubber toughening of polypropylene continues to be an active area of investigation.[69] It is well known from previous studies that, among other factors, the rubber particle diameter is of critical importance.[69–71] These studies indicate that the modification in the rubber domain size to an optimum diameter improves the impact strength of PP. It is also well known that the addition of the rubber to polypropylene results in a decrease in stiffness.[72] Thus although successes in improving the impact properties of PP have been attained by rubber toughening, its application, e.g., in the automotive industry, is still limited because of the opposing trends of stiffness and toughness.

An attractive strategy to improve the stiffness, while maintaining the impact strength of PP and rubber-modified PP, is the formation of nanocomposites. Addition of mineral fillers, such as talc, $CaCO_3$, and glass to PP, provides an increase in stiffness, but these particulates are stress concentrators and result in concomitant decrease in impact strength.[73,74] The fact that the filler is at the nanometer-size scale in these systems gives promise of achieving improved stiffness without significant loss of impact strength.[75] This is under the assumption that the clay platelets do not act as stress concentrators because of their nanometer size scale. Thus this desirable combination of improved stiffness and maintenance of impact strength by the addition of clay in TPOs, if attained, could boost the application of TPOs in the automotive and other industries.

For example, Oldenbo[76] from Volvo Car Corp., Sweden, recently presented data regarding the effect of addition of clay on the impact and flexural properties of TPO/clay nanocomposites. The desirable combination of improvement in stiffness and maintenance of impact

strength was not observed in that study. Monotonic increase in stiffness was observed, but a precipitous decrease in impact was observed, as clay loading increased. The decrease in impact strength was ascribed to the lack of good dispersion of the clay. The existence of clay tactoids and "chunks" was supposed to be the cause of poor impact strength because these acted as stress concentrators.

Very little work has been published on the development of TPO nanocomposites. Much of what has been published resulted from the joint development of TPO nanocomposites for injection-molded automotive components by General Motors Research and Basell Polyolefins. This joint development was described in a publication, which highlighted the process ability advantages of a TPO/clay nanocomposite over a conventional TPO/talc composite.[77] This TPO nanocomposite is covered in patents assigned to Basell Poliolefine Italia S.p.A.[78] and Basell Technology Company BV.[79] No reports on the morphology and the specific degree of intercalation and exfoliation of the clay in the material developed by General Motors Research and Basell Polyolefins have been published.

This development and other similar developments of TPO/clay nanocomposites for automotive and other applications were heavily reported in the trade literature.[80–87] In one of these trade journal articles, it was claimed that, "much of the development work done before 1998 with natural montmorillonites having amorphous silica impurities is of little practical value."[83] This claim was based on the proposition that large proportions of amorphous silica in the clay substantially degrade impact strength and elongation, increase opacity or haze levels, and increase gas permeability.

A study of the requirements for preparing exfoliated TPO/C_{18}-mmt nanocomposites was reported.[88] It was noted that nanocomposite properties were dependent not only on exfoliation of the clay, but also on the structure and properties of the silicate / polymer matrix interface. The optimization of the organoclay structure for the maximum enhancement of TPO/clay nanocomposite mechanical properties was discussed.

The effect of varying clay loading on the morphology of TPO/clay nanocomposites was studied using atomic force microscopy, transmission electron microscopy, and X-ray diffraction.[89] Thermoplastic olefin/clay nanocomposites were made at clay loadings varying from 0.6 to 6.7 wt.%. The morphology of these TPO/clay nanocomposites was investigated with atomic force microscopy (AFM), transmission electron microscopy (TEM), and X-ray diffraction (XRD). The ethylene-propylene rubber (EPR) particle morphology in the TPO was found to undergo progressive particle breakup and decrease in particle size, as clay loading increased in the range from 0.6 to 5.6 wt.% clay. Transmission electron micrographs showed that the clay platelets preferentially segregate to the rubber–particle interface. The breakup of the EPR particles was suspected to be a result of the increasing melt viscosity observed as clay loading increased and/or the accompanying chemical modifiers on the clay, acting as interfacial agents, reducing the interfacial tension with concomitant reduction in particle size. The clay morphology was mainly as intercalated tactoids with a minor amount of exfoliated and dispersed clay platelets evident in the TEM photomicrographs. Flexural modulus increased monotonically as clay loading increased. Impact strength was maintained or decreased modestly as clay loading increased.

CONCLUSION

Polypropylene is one of the most widely used thermoplastics not only because of its balance of physical and mechanical properties, but also because of its environmental friendliness (e.g., recyclability) and low cost. Although the addition of mineral fillers to polypropylenes enhances the stiffness properties, the concomitant effect is typically a decrease in toughness properties because the inorganic filler particles act as stress concentrators. Therefore the formation of propylene-based polymer/clay nanocomposites is an area of intensive investigation.

Considering that clay is composed of platelets that are about 1 nm in thickness and around 1000 nm in lateral dimensions and assuming a density of 2.65 g/cm^3 yields an upper limit of about 750 m^2/g surface area per gram for fully exfoliated clay. This is 5 to 10 times the surface area per gram achievable with ordinary mineral fillers. Exfoliated clay in polymers has been shown to endow unique properties to the system, such as increased Tg, increased stiffness, maintenance of impact strength, increased barrier to gases and liquids, increased flame retardancy, etc.

The tightly bound stacks of clay platelets in natural clay may be readily separated by simple dissolution in a polar solvent; however, separation of the clay platelets in a nonpolar polymer is difficult. Therefore organic modification of natural clay minerals is required to disperse the clay in polypropylenes. This typically involves exchange of cations on the natural clay with alkyl ammonium cations with alkyl chains of varying lengths. However, the vast majority of experimental studies and commercial developments employ polypropylene with maleated polypropylene (PP-MA), or more rarely another functionalized PP, as a compatibilizer to form PP/clay nanocomposites with reasonable dispersion of the clay.

These polypropylene nanocomposites have exhibited enhancements in mechanical properties, rheological properties, barrier to gases and liquids, flammability, thermal

stability, heat deflection temperature, and electrical properties. Although the majority of work that has been published, relative to propylene-based polymers, concerns the development of polypropylene nanocomposites, there has been little incentive to commercially produce these compositions because polypropylene is a very low cost commodity thermoplastic and the improvements in the properties so far demonstrated have not justified the increased cost to produce nanocomposites commercially.

On the other hand, far more interest has been generated in developing thermoplastic olefin (TPO) nanocomposites. The applications of TPOs, especially targeted for nanocomposite development, are those in the automotive sector. The major initiative in the development of TPO nanocomposites was that of General Motors Research and Basell Polyolefins to develop TPO nanocomposites for injection-molded automotive components. Automotive TPOs are usually reinforced with large quantities (up to 50 vol.%) of mineral fillers, such as talc and calcium carbonate. Replacing these conventional composites with lighter weight and better-performing TPO nanocomposites has driven these efforts. Because of the much higher cost basis of TPOs, relative to polypropylene homopolymer, this objective appears to be much easier to achieve. The processability advantages of a TPO/clay nanocomposite over a conventional TPO/talc composite have been reported, but no reports on the morphology and the specific degree of intercalation and exfoliation of the clay in the material developed by General Motors Research and Basell Polyolefins have been published.

Therefore the state of development of polypropylene homopolymer or TPO nanocomposites is currently in an early stage. From the technical standpoint, the formation of nanocomposites with uniformly dispersed clay remains a daunting challenge. From the commercial standpoint, the production of economically viable nanocomposites, i.e., with properties enhanced to the extent that the increased cost of production is justified, also remains a daunting challenge. Certainly, the research that is progressively reported, which better defines and innovates upon the structure and properties of polypropylene and TPO nanocomposites, continues to spur interest in commercial development. The outlook remains bright, but formidable challenges, both technical and commercial, abound.

REFERENCES

1. Kojima, Y.; Usuki, A.; Kawasumi, M.; Okada, A.; Fukushima, Y.; Kurauchi, T.; Kamigaito, O. Mechanical properties of nylon 6–clay hybrid. J. Mater. Res. **1993**, *8* (5), 1185–1189.
2. Kojima, Y.; Usuki, A.; Kawasumi, M.; Okada, A.; Kurauchi, T.; Kamigaito, O. Synthesis of nylon-6–clay hybrid by montmorillonite intercalated with epsilon-caprolactam. J. Polym. Sci., Part A, Polym. Chem. **1993**, *31* (4), 983–986.
3. Usuki, A.; Kojima, Y.; Kawasumi, M.; Okada, A.; Fukushima, Y.; Kurauchi, T.; Kamigaito, O. Synthesis of nylon 6–clay hybrid. J. Mater. Res. **1993**, *8* (5), 1179–1184.
4. Vaia, R.A.; Ishii, H.; Giannelis, E.P. Polymer-layered silicate nanocomposites: In situ intercalative polymerization of epsilon-caprolactone in layered silicates. Chem. Mater. **1993**, *5* (8), 1064–1066.
5. Manias, E.; Chen, H.; Krishnamoorti, R.; Genzer, J.; Kramer, E.J.; Giannelis, E.P. Intercalation kinetics of long polymers in 2 nm confinements. Macromolecules **2000**, *33* (21), 7955–7966.
6. Manias, E.; Touny, A.; Wu, L.; Strawhecker, K.; Lu, B.; Chung, T.C. Polypropylene/montmorillonite nanocomposites. Review of the synthetic routes and materials properties. Chem. Mater. **2001**, *13* (10), 3516–3523.
7. Wilson, P.S. Nanocomposite Market Opportunities. In *Nanocomposites 2001: Delivering New Value to Plastics*, Executive Conference Management, Chicago, June 25–27, 2001; Plymouth, MI, 2001.
8. Sinclair, K. The Market Potential for Nanocomposites: Quantifying the Opportunities. In *Nanocomposites 2001: Delivering New Value to Plastics*, Executive Conference Management, Chicago, June 25–27, 2001; Plymouth, MI, 2001.
9. Kato, M.; Usuki, A.; Okada, A. Synthesis of polypropylene oligomer–clay intercalation compounds. J. Appl. Polym. Sci. **1997**, *66* (9), 1781–1785.
10. Usuki, A.; Kato, M.; Okada, A.; Kurauchi, T. Synthesis of polypropylene–clay hybrid. J. Appl. Polym. Sci. **1997**, *63* (1), 137–139.
11. Kawasumi, M.; Hasegawa, N.; Kato, M.; Usuki, A.; Okada, A. Preparation and mechanical properties of polypropylene–clay hybrids. Macromolecules **1997**, *30* (20), 6333–6338.
12. Xu, W.; Liang, G.; Wang, W.; Tang, S.; He, P.; Pan, W.-P. PP-PP-g-MA-Org-MMT nanocomposites. I. Intercalation behavior and microstructure. J. Appl. Polym. Sci. **2003**, *88* (14), 3225–3231.
13. Walter, P.; Mader, D.; Reichert, P.; Mulhaupt, R.J. Novel polypropylene materials. J. Macromol. Sci., Pure Appl. Chem. **1999**, *A36*, 1613–1639.
14. Reichert, P.; Nitz, H.; Klinke, S.; Brandsch, R.; Thomann, R.; Mulhaupt, R. Poly(propylene)/organoclay nanocomposite formation: Influence of compatibilizer functionality and organoclay modification. Macromol. Mater. Eng. **2000**, *275* (2), 8–17.

15. Hofmann, U. The chemistry of clays possessing a high swelling capacity (bentonites). Angew. Chem. **1956**, *68*, 53–61.
16. Weiss, A. Organic derivatives of micaceous layered silicates. Angew. Chem. **1963**, *75*, 113–122.
17. Reichert, P.; Hoffmann, B.; Bock, T.; Thomann, R.; Mulhaupt, R.; Friedrich, C. Morphological stability of poly(propylene) nanocomposites. Macromol. Rapid Commun. **2001**, *22* (7), 519–523.
18. Southern Clay Products Inc. *Material Safety Data Sheet, Cloisite 15A and 20A*.
19. Morgan, A.B.; Gilman, J.W. Characterization of polymer-layered silicate (clay) nanocomposites by transmission electron microscopy and X-ray diffraction: A comparative study. J. Appl. Polym. Sci. **2003**, *87* (8), 1329–1338.
20. Kim, K.-N.; Kim, H.; Lee, J.-W. Effect of interlayer structure, matrix viscosity and composition of a functionalized polymer on the phase structure of polypropylene–montmorillonite nanocomposites. Polym. Eng. Sci. **2001**, *41* (11), 1963–1969.
21. Zhang, Q.; Wang, Y.; Fu, Q. Shear-induced change of exfoliation and orientation in polypropylene/montmorillonite nanocomposites. J. Polym. Sci., Part B, Polym. Phys. **2003**, *41* (1), 1–10.
22. Chen, L.; Wong, S.-C.; Pisharath, S. Fracture properties of nanoclay-filled polypropylene. J. Appl. Polym. Sci. **2003**, *88* (14), 3298–3305.
23. Liu, X.; Wu, Q. PP/clay nanocomposites prepared by grafting-melt intercalation. Polymer **2001**, *42* (25), 10013–10019.
24. Kaempfer, D.; Thomann, R.; Mulhaupt, R. Melt compounding of syndiotactic polypropylene nanocomposites containing organophilic layered silicates and in situ formed core/shell nanoparticles. Polymer **2002**, *43* (10), 2909–2916.
25. Koo, C.M.; Kim, S.O.; Chung, I.J. Study on morphology evolution, orientational behavior, and anisotropic phase formation of highly filled polymer-layered silicate nanocomposites. Macromolecules **2003**, *36* (8), 2748–2757.
26. Somwangthanaroj, A.; Lee, E.C.; Solomon, M.J. Early stage quiescent and flow-induced crystallization of intercalated polypropylene nanocomposites by time-resolved light scattering. Macromolecules **2003**, *36* (7), 2333–2342.
27. Marchant, D.; Jayaraman, K. Strategies for optimizing polypropylene–clay nanocomposite structure. Ind. Eng. Chem. Res. **2002**, *41* (25), 6402–6408.
28. Ma, J.; Qi, Z.; Hu, Y. Synthesis and characterization of polypropylene/clay nanocomposites. J. Appl. Polym. Sci. **2001**, *82* (14), 3611–3617.
29. Vaia, R.A.; Giannelis, E.P. Lattice model of polymer melt intercalation in organically-modified layered silicates. Macromolecules **1997**, *30* (25), 7990–7999.
30. Vaia, R.A.; Giannelis, E.P. Polymer melt intercalation in organically-modified layered silicates: Model predictions and experiment. Macromolecules **1997**, *30* (25), 8000–8009.
31. Fornes, T.D.; Yoon, P.J.; Keskkula, H.; Paul, D.R. Nylon 6 nanocomposites: The effect of matrix molecular weight. Polymer **2001**, *42* (25), 9929–9940.
32. Schon, F.; Thomann, R.; Gronski, W. Shear controlled morphology of rubber/organoclay nanocomposites and dynamic mechanical analysis. Macromol. Symp. **2002**, *189* (1), 105–110.
33. Xu, W.; Ge, M.; He, P. Nonisothermal crystallization kinetics of polypropylene/montmorillonite nanocomposites. J. Polym. Sci., Part B, Polym. Phys. **2002**, *40* (5), 408–414.
34. Hambir, S.; Bulakh, N.; Jog, J.P. Polypropylene/clay nanocomposites: Effect of compatibilizer on the thermal, crystallization and dynamic mechanical behavior. Polym. Eng. Sci. **2002**, *42* (9), 1800–1807.
35. Pozsgay, A.; Frater, T.; Papp, L.; Sajo, I.; Pukanszky, B. Nucleating effect of montmorillonite nanoparticles in polypropylene. J. Macromol. Sci., Phys. **2002**, *B41* (4–6), 1249–1265.
36. Maiti, P.; Nam, P.H.; Okamoto, M.; Hasegawa, N.; Usuki, A. Influence of crystallization on intercalation, morphology, and mechanical properties of polypropylene/clay nanocomposites. Macromolecules **2002**, *35* (6), 2042–2049.
37. Nam, P.H.; Maiti, P.; Okamoto, M.; Kotaka, T.; Hasegawa, N.; Usuki, A. A hierarchical structure and properties of intercalated polypropylene/clay nanocomposites. Polymer **2001**, *42* (23), 9633–9640.
38. Hambir, S.; Bulakh, N.; Kodgire. P.; Kalgaonkar, R.; Jog, J.P. PP/clay nanocomposites: A study of crystallization and dynamic mechanical behavior. J. Polym. Sci., Part B, Polym. Phys. **2001**, *39* (4), 446–450.
39. Maiti, P.; Nam, P.H.; Okamoto, M.; Kotaka, T.; Hasegawa, N.; Usuki, A. The effect of crystallization on the structure and morphology of polypropylene/clay nanocomposites. Polym. Eng. Sci. **2002**, *42* (9), 1864–1871.
40. Svehlova, V.; Poloucek, E. Mechanical-properties of talc-filled polypropylene—Influence of filler content, filler particle-size and quality of dispersion. Angew. Makromol. Chem. **1994**, *214*, 91–99.
41. Maiti, S.N.; Mahapatro, P.K. Mechanical-properties of I-PP/CaCO3 composites. J. Appl. Polym. Sci. **1991**, *42* (12), 3101–3110.
42. Hasegawa, N.; Kawasumi, M.; Kato, M.; Usuki, A.; Okada, A. Preparation and mechanical properties of

polypropylene–clay hybrids using a maleic anhydride-modified polypropylene oligomer. J. Appl. Polym. Sci. **1998**, *67* (1), 87–92.
43. Hasegawa, N.; Okamoto, H.; Kato, M.; Usuki, A. Preparation and mechanical properties of polypropylene–clay hybrids based on modified polypropylene and organophilic clay. J. Appl. Polym. Sci. **2000**, *78* (11), 1918–1922.
44. Zhang, Q.; Wang, K.; Men, Y.; Fu, Q. Dispersion and tensile behavior of polypropylene/montmorillonite nanocomposites produced via melt intercalation. Chin. J. Polym. Sci. **2003**, *21* (3), 359–367.
45. Kodgire, P.; Kalgaonkar, R.; Hambir, S.; Bulakh, N.; Jog, J.P. PP/clay nanocomposites: Effect of clay treatment on morphology and dynamic mechanical properties. J. Appl. Polym. Sci. **2001**, *81* (7), 1786–1792.
46. Hernandez-Luna, A.; D'Souza, N.A.; Renade, A.; Drewniak, M. Skin-core Effects in Polypropylene Nanocomposites. In *Affordable Materials Technology: Platform to Global Value and Performance*, 47th International SAMPE Symposium, Long Beach, CA, May 12–16 2002; Society for the Advancement of Material and Process Engineering: Covina, CA 2002; 1100–1106.
47. Wong, S.-C.; Chen, L. Mechanical and Fracture Properties of Nanoclay-Filled Polypropylene. In *ANTEC 2002, San Francisco*, CA, May 5–9, 2002; Society of Plastics Engineers: Brookfield, CT, 2002; Vol. 2, 1466–1469.
48. Svoboda, P.; Zeng, C.; Wang, H.; Lee, L.J.; Tomasko, D.L. Morphology and mechanical properties of polypropylene/organoclay nanocomposites. J. Appl. Polym. Sci. **2002**, *85* (7), 1562–1570.
49. Rong, M.Z.; Zhang, M.Q.; Zheng, Y.X.; Zeng, H.M.; Walter, R.; Friedrich, K. Structure–property relationships of irradiation grafted nano-inorganic particle filled polypropylene composites. Polymer **2001**, *42* (1), 167–183.
50. Brune, D.A.; Biccerano, J. Micromechanics of nanocomposites: Comparison of tensile and compressive elastic moduli, and prediction of effects of incomplete exfoliation and imperfect alignment on modulus. Polymer **2002**, *43* (2), 369–387.
51. Solomon, M.J.; Abdulwahab, A.S.; Seefeldt, K.F.; Somwangthanaroj, A.; Varadan, P. Rheology of polypropylene/clay hybrid materials. Macromolecules **2001**, *34* (6), 1864–1872.
52. Reichert, P.; Bock, T.; Thomann, R.; Mulhaupt, R.; Friedrich, C. Morphological stability of poly(propylene) nanocomposites. Macromol. Rapid Commun. **2001**, *22* (7), 519–523.
53. Marchant, D.; Jayaraman, K. Rheological Probing of Structure in Polypropylene/Clay Nanocomposites. In *ANTEC 2001,* Dallas, TX, May 6–10, 2001; Society of Plastics Engineers: Brookfield, CT, 2001; Vol. 2, 2130–2134.
54. Krishnamoorti, R.; Ren, J.; Siva, A.S. Shear response of layered silicate nanocomposites. J. Chem. Phys. **2001**, *114* (11), 4968–4973.
55. Nam, P.H.; Maiti, P.; Okamoto, M.; Kotaka, T.; Nakayama, T.; Takada, M.; Oshima, M. Foam processing and cellular structure of polypropylene/clay nanocomposites. Polym. Eng. Sci. **2002**, *42* (9), 1907–1918.
56. Koo, C.M.; Kim, M.J.; Choi, M.H.; Kim, S.O.; Chung, I.J. Mechanical and rheological properties of the maleated polypropylene-layered silicate nanocomposites with different morphology. J. Appl. Polym. Sci. **2003**, *88* (6), 1526–1535.
57. Boucard, S.; Duchet, J.; Gerard, J.F.; Prele, P.; Gonzalez, S. Processing of polypropylene–clay hybrids. Macromol. Symp. **2003**, *194* (1), 241–246.
58. Galgali, G.; Ramesh, C.; Lele, A. A rheological study on the kinetics of hybrid formation in polypropylene nanocomposites. Macromolecules **2001**, *34* (4), 852–858.
59. Lele, A.; Mackley, M.; Galgali, G.; Ramesh, C. In situ rheo-X-ray investigation of flow-induced orientation in layered silicate-syndiotactic polypropylene nanocomposite melt. J. Rheol. **2002**, *46* (5), 1091–1110.
60. Mehta, S.D.; Shankernarayanan, M.J.; Mavridis, H. *PCT Application 03/022917 A1*; March 20 2003.
61. Kenig, S.; Ophir, A.; Shepelev, O.; Weiner, F. High Barrier Blow Molded Containers Based on Nano Clay Composites. In *ANTEC 2002, San Francisco,* CA, May 5–9, 2002; Society of Plastics Engineers: Brookfield, CT, 2002; Vol. 2, 794–798.
62. Manias, E.; Touny, A.; Wu, L.; Lu, B.; Chung, T.C. Polypropylene/silicate nanocomposites, synthetic routes and materials properties. Polym. Mater. Sci. Eng. **2000**, *82*, 282–283.
63. Gilman, J.W.; Jackson, C.L.; Morgan, A.B.; Harris, R., Jr.; Manias, E.; Giannelis, E.P.; Wuthenow, M.; Hilton, D.; Phillips, S.H. Flammability properties of polymer-layered-silicate nanocomposites. Polypropylene and polystyrene nanocomposites. Chem. Mater. **2000**, *12* (7), 1866–1873.
64. Wagenknecht, U.; Kretzschmar, B.; Reinhardt, G. Investigations of fire retardant properties of polypropylene–clay-nanocomposites. Macromol. Symp. **2003**, *194* (1), 207–212.
65. Wang, D.; Wilkie, C.A. In-situ reactive blending to prepare polystyrene–clay and polypropylene–clay nanocomposites. Polym. Degrad. Stab. **2003**, *80* (1), 171–182.
66. Morgan, A.B.; Harris, J.D. Effects of organoclay

Soxhlet extraction on mechanical properties, flammability properties and organoclay dispersion of polypropylene nanocomposites. Polymer **2003**, *44* (8), 2313–2320.
67. Zanetti, M.; Camino, G.; Reichert, P.; Mulhaupt, R. Thermal behaviour of poly(propylene) layered silicate nanocomposites. Macromol. Rapid Commun. **2001**, *22* (3), 176–180.
68. Shen, J.-W.; Chen, X.M.; Huang, W.-Y. Structure and electrical properties of grafted polypropylene/graphite nanocomposites prepared by solution intercalation. J. Appl. Polym. Sci. **2003**, *88* (7), 1864–1869.
69. Jang, B.Z.; Uhlmann, D.R.; Vander Sande, J.B. Rubber-toughening in polypropylene. J. Appl. Polym. Sci. **1985**, *30* (6), 2485–2504.
70. Jang, B.Z.; Uhlmann, D.R.; Vander Sande, J.B. The Rubber Particle Size Dependence of Crazing in Polypropylene. In *ANTEC 84,* New Orleans, LA, April 30–May 3, 1984; Society of Plastics Engineers: Brookfield, CT, 1984; 549–551.
71. Mehrabzadeh, M.; Hossein Nia, K. Impact modification of polypropylene by ethylene propylene copolymer-grafted maleic anhydride. J. Appl. Polym. Sci. **1999**, *72* (10), 1257–1265.
72. Lu, J.; Wei, G.-X.; Sue, H.-J.; Chu, J. Toughening mechanisms in commercial thermoplastic polyolefin blends. J. Appl. Polym. Sci. **2000**, *76* (3), 311–319.
73. Radosta, J.A. Improving the Impact-Stiffness Balance of Mineral Filled Polyolefins with Surface Modified Talc. In *ANTEC 84,* New Orleans, LA, April 30–May 3, 1984; Society of Plastics Engineers: Brookfield, CT, 1984; 145–149.
74. Marshall, C.J.; Rozett, R.; Kunkle, A.C. Effects of mica as a filler in polyolefins and PVC. Plast. Compd. **1985**, *8* (7), 69–74.
75. Svoboda, P.; Zeng, C.; Wang, H.; Lee, L. Structure and Mechanical Properties of Polypropylene/Organoclay Nanocomposites. In *Nanocomposites 2001: Delivering New Value to Plastics*; Executive Conference Management, Plymouth, MI. 2001.
76. Oldenbo, M. TPO-Nanocomposites for Automotive Exterior Bodypanels: Potentials and Experiences from Evaluations of Commercial Materials. In *Nanocomposites 2001: Delivering New Value to Plastics*; Executive Conference Management: Chicago, June 25–27, Plymouth, MI, 2001.
77. Okonski, D.A. Injection molding a polyolefin-based nanocomposite versus a talc-filled TPO. Mater. Res. Symp. Proc. **2002**, *702*, 265–276.
78. Bishop, C.E.; Niyogi, S.G. Intercalated Clay Useful for Making an Alpha-Olefin Polymer Material Nanocomposite. US Patent 6,500,892, December 31, 2002.
79. Niyogi, S.G. Nanocomposites Made from Polypropylene Graft Copolymers. US Patent 6,451,897, September 17, 2002.
80. US polymer manufacturers are working on nanocomposites. Addit. Polym. **1999**, *1999* (2), 9–10.
81. Grande, J.A. Olefin-based nanocomposites hold potential for automotive. Mod. Plast. Int. **1999**, *29* (1), 35–36.
82. Garces, J.M.; Moll, D.J.; Bicerano, J.; Fibiger, R.; McLeod, D. Polymeric nanocomposites for automotive applications. Adv. Mater. **2000**, *12* (23), 1835–1839.
83. Mapleston, P. Tiny mineral fillers bring big benefits in compounding. Mod. Plast. Int. **2000**, *30* (11), 72–74.
84. Murphy, J. Nanocomposites—It's a question of picking the winners. Plast. Addit. Compound. **2000**, *2* (5), 30–32.
85. Leaversuch, R. Nanocomposites: Broaden roles in automotive, barrier packaging. Plast. Technol. **2001**, *47* (10), 64–69.
86. Nanocomposite body panels. High Perform. Plast. **2001**, 8–9.
87. Sall, K. Smaller is better. Eur. Plast. News **2002**, *29* (3), 14.
88. Merinska, D.; Malac, Z.; Hrncirik, J.: Simonik, J.; Trlica, J.; Pospisil, M.; Capkova, P.; Weiss, Z. Modification of Clay Intercalate Structure and Properties of TPO Based Nanocomposites. In *ANTEC 2001,* Dallas, TX, May 6–10, 2001; Society of Plastics Engineers: Brookfield, CT, 2001; Vol. 2, 2166–2170.
89. Mehta, S.; Mirabella, F.M.; Rufener, K.; Bafna, A. TPO/clay nanocomposites: Morphology and mechanical properties. J. Appl. Polym. Sci. *in press.*

Protein Adsorption Kinetics Under an Applied Electric Field

Paul R. Van Tassel
Yale University, New Haven, Connecticut, U.S.A.

INTRODUCTION

The dimensions of the largest naturally occurring molecular species and those of the smallest manmade features converge at the nanoscale. Proteins, the basic building blocks of living organisms, are among the larger molecules in nature with dimensions ranging from 1 to 100 nm. Materials incorporating proteins are therefore true nanomaterials. Of particular importance are thin films of proteins immobilized at a solid substrate. Biosensing, tissue engineering, enzymatic catalysis, and bioelectronics are just a few of the areas in which immobilized layers of proteins play a key role.

The tendency of proteins to attach to interfacial regions is well documented.[1–4] Ionic, van der Waals, solvation, and donor–acceptor interactions all play important roles in rendering the interfacially adsorbed state to be thermodynamically favored over the solution state.[1] Proteins are colloidal objects, possessing a distribution of surface charge and, in an electrolytic solution, a distribution of weakly associated counterions. Their interaction with a solid substrate is thus expected to be sensitive to the substrate's charge distribution. By controlling the polarization of an adsorbing surface (i.e., applying an electric field), one alters this charge distribution and therefore the surface–protein interaction. This possibility is understandably appealing to those wishing to control the adsorption process, perhaps desiring adsorbed layers of preferred orientation or spatial distribution. However, the interaction between proteins and surfaces is complex and predicting adsorbed layer properties by considering the contributions from the interaction modes listed above remains a significant challenge. Adding an electric field makes the problem even more complex. Thus while influencing an adsorbed protein layer with an electric field is both possible and desirable, the outcome is as yet poorly understood.

In this contribution, we review the field of protein adsorption kinetics under an applied electric field. By focusing on kinetics, we limit ourselves to studies where adsorbed layer properties are measured in situ during the electroformation process. We begin with a brief presentation of certain basic theoretical considerations. We then introduce the methods employed to measure protein adsorption kinetics under an applied electric field. Next, we introduce some of the key results, grouping our presentation by investigator. A perspective on future directions is then given and this is followed by a conclusion. By summarizing some of the key accomplishments and open questions, we hope to guide future efforts to produce nanoscale devices employing adsorbed protein layers formed under an electric field.

THEORY

An applied electric field may affect protein adsorption in two ways. The first of these deals with transport. Proteins are composed of amino acids, some of which contain acidic/basic sites. Thus at all but the isoelectric pH, the protein will possess a net charge and thus migrate in response to an electric field. Because the charge distribution is generally not spherically symmetric, the electric field also imposes a torque on the molecule, causing it to rotate. Of course, electric-field-induced migration and rotation must compete against the molecule's thermal diffusive motion, so these influences are only observed in excess of some threshold field strength. The second principal effect of an applied electric field involves the direct interaction between the electrode and the protein upon close approach. The alteration of electrode potential brings about various physicochemical changes in the surface itself and alters the counterion distribution away from the surface; these two effects influence the protein–surface interaction and affect such properties as the rate of attachment, protein orientation, and the rate and extent of subsequent post-adsorption changes in orientation and conformation.

The transport of protein in a (assumedly laminar) flowing solution to the surface may be described by the general transport relation

$$\frac{\partial c}{\partial t} + \vec{v} \cdot \vec{\nabla} c = \vec{\nabla} \cdot \left(D \vec{\nabla} c + \frac{cq\vec{E}}{\zeta} \right) \tag{1}$$

where c is the protein concentration, t is the time, $\vec{v}$ is the fluid flow velocity vector, D is the diffusivity, q is the effective protein charge, $\vec{E}$ is the electric field vector, and ζ is the friction coefficient (a measure of the viscous drag

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120014126

on the protein). The flow field must be known in advance to solve Eq. 1. So long as the concentration is quite low, one may neglect the influence of the migrating proteins on the fluid and solve for the flow field using the standard Navier–Stokes equation.

The two terms on the right side of Eq. 1 represent the contributions from thermal diffusion and electric-field-induced migration, respectively. It is interesting to consider the limiting cases where one is much larger than the other. Neglecting the second term gives pure convective diffusion. One may approximate the flux to the surface by assuming steady shear flow and a perfect sink boundary condition at the adsorbing surface. In this case, the flux may be approximated by $J(x, t)=Dc_b/\delta(x, t)$, where c_b is the bulk protein concentration and

$$\delta(x,t) = \frac{(45Dx/2a)^{1/3}}{f\left(\dfrac{4Dt}{(45Dx/2a)^{2/3}}\right)} \tag{2}$$

is the convective diffusion boundary layer, a function of the diffusivity (D), the distance along the surface in the direction of the flow (x), the shear rate (a), and the time (t).[5,6] The function $f(\tau)$ is the inverse of $\tau(f)=[1-(1-f^3)^{2/3}]/2f^2$ for $\tau<1/2$ and $f(\tau)=1$ for $\tau>1/2$. Pure convective diffusion is therefore characterized by an initially zero flux that increases steadily to a steady flux at dimensionless time, τ, equals 1/2.

Neglecting the first term in Eq. 1 is equivalent to ignoring diffusive motion. In this case, assuming an electric field normal to the surface of magnitude E, fully developed shear flow, and a no-slip, perfect sink boundary condition, the concentration profile is given by

$$c(x,z,t) = c_b H\left(azt + \frac{aqEt^2}{2\zeta} - x\right) \tag{3}$$

where z is the height above the surface and H is the Heaviside function [i.e., $H(\xi)=1$ for $\xi>0$ and $H(\xi)=0$ for $\xi<0$].[6] The flux to the surface is thus

$$J(x,t) = \frac{c_b qE}{\zeta} H\left(\frac{aqEt^2}{2\zeta x} - 1\right) \tag{4}$$

Pure electrophoretic migration is therefore characterized by an initial period of zero flux, during a time of $\sqrt{2\zeta x/aqE}$, followed by a steady flux.

In the absence of transport limitations, the rate of adsorption may be expressed as

$$\frac{d\Gamma}{dt} = k_a c_b \Phi - \sum_i k_{d,i}\Gamma_i \tag{5}$$

where Γ is the density of adsorbed protein (mass per area), k_a is the adsorption rate constant, Φ is the one-body cavity function, and $k_{d,i}$ and Γ_i are the desorption rate constant and the density of protein in the ith structural state, respectively (these states may denote various conformations, orientations, or states of aggregation).[7] The cavity function is defined as $\Phi=\langle e^{-u/kT}\rangle_{\Gamma,T}$, where u is the potential energy of a single molecule interacting with the surface and with all of the previously adsorbed molecules (u depends on position and orientation), k is the Boltzmann constant, T is the absolute temperature, and the brackets represent an averaged quantity, over all representations of the adsorbed layer at density Γ and temperature T, according to their appropriate weights, and over all orientations and positions of the single "reference" molecule. All of the quantities on the right of Eq. 5 (except c_b) may be altered by application of an electric field.

To make quantitative predictions, the potential energy of interaction between protein molecules and the charged surface must be calculated. The electrostatic contribution to this energy may be determined using Poisson–Boltzmann[8–12] or density functional[13,14] approaches.

EXPERIMENTAL METHODS

In this section, we review the experimental methods used to measure protein adsorption kinetics, in situ, under an applied electric field.

Impedance Measurements

Impedance measurements involve the application of an alternating voltage across an electrode / solution interface.[15–17] The solution is modeled as a resistor and the interface as an impeder and a capacitor in series. The capacitance is simply the imaginary part of the measured impedance. Adsorbed molecules (e.g., organics and proteins) displace ions and solvent from the interface; this changes the charge distribution and thereby the interfacial capacitance. The capacitance is thus an indirect measure of the extent of adsorption and structure of the adsorbed layer.

Reflectometry

Reflectometry involves measuring the reflection of polarized light at the interface between two optical media. In fact, two reflections are measured: one for the electric field component perpendicular to the plane of incidence (transverse electric or s-wave) and one for the electric field component parallel to the plane of incidence (transverse magnetic or p-wave). At a certain angle of incidence (the Brewster angle), the p-wave reflection vanishes and around this angle, the reflectivity, or square of the amplitude of the p-wave reflection, and ellipticity, or ratio of p- and s-wave reflections, become very

sensitive to interfacial heterogeneity, as brought about, for example, by adsorption of biomolecules. By assuming the adsorbed layer to be uniform in refractive index, both its thickness and refractive index may be determined. By further assuming a linear dependence of refractive index on concentration, the adsorbed density is calculable. Silica is typically used as the adsorbing substrate; thus to apply an electric field, a conducting coating is needed. Bos et al.[18] have developed a reflectometry method using silica coated with a ca. 100-nm layer of indium tin oxide (ITO). An important caveat is that the intensity ratio changes in the presence of an applied voltage and this must be considered when calculating the amount of adsorbed protein.

Quartz Crystal Microbalance

A quartz crystal microbalance (QCM) consists of a thin disk of (piezoelectric) crystalline quartz sandwiched between thin film metal electrodes. Upon application of an alternating voltage, the crystal undergoes thickness shear mode vibration. The mass adsorbed to the electrode surface, including the trapped solvent, is simply proportional to the frequency shift. In addition, the dissipation of energy following voltage removal, as measured by the decay of the oscillation amplitude, is a sensitive measure of the viscoelastic properties of an adsorbed layer. Khan and Wernet[19] have developed a QCM method in which a voltage is applied between an electroconductive polymer film-coated adsorbing surface and a counterelectrode.

Optical Waveguide Lightmode Spectroscopy

Optical waveguide lightmode spectroscopy (OWLS) is based on the phase shift associated with multiple interfacial reflections: when either the s- or p-wave undergoes a total phase shift equal to an integral multiple of 2π upon one complete traversal of a planar, dielectric waveguide sandwiched between media of lower refractive index, a standing wave is excited in the waveguiding film. Because of their dependence on reflection, the phase shifts are sensitive to interfacial heterogeneity—and the thickness, refractive index, and density of an adsorbed biomolecular layer can be readily determined. Optical waveguide lightmode spectroscopy is normally applied to nonconductive substrates. To allow for detection in the presence of an electric field, Brusatori et al.[6,20] and Bearinger et al.[21] have developed OWLS methods employing an indium tin oxide coating.

Total Internal Reflection Fluorescence

When light traversing an optically dense medium approaches an interface with a more optically rare medium at an angle exceeding a critical value of $\theta_{crit}=\sin^{-1}(n_{rare}/n_{dens})$, a total internal reflection occurs and an evanescent wave of exponentially decaying intensity penetrates the rarer medium. In total internal reflection fluorescence (TIRF) spectroscopy, the evanescent wave excites fluorescent probes attached to the biomolecules and detection of the emission associated with their decay provides information on the density, composition, and conformation of adsorbed molecules. Fraaije et al.[22] and Asanov et al.[23,24] have extended TIRF methods to measure protein adsorption under an applied electric field.

RESULTS

Bernabeu and Caprani[15] have measured the adsorption of fibrinogen and albumin to a carbon electrode using an impedance method. They observed the initial rate and the overall extent of adsorption of these two negatively charged proteins to decrease with increasing voltage, in apparent violation of electrostatic rules Fig. 1. The authors explained this in terms of adsorbed counterions. On a platinum (i.e., more hydrophilic) surface, the rate and overall extent yielded minima at an intermediate electrode potential.

Feng and Andrade[16] have investigated the adsorption of several proteins on low-temperature isotropic carbon electrodes, also using an impedance method. They generally find the extent and initial rate of adsorption to decrease with increasing electrode potential, irrespective of protein charge and solution ionic strength Fig. 2. In contrast, on a gold electrode, the extent and initial rate varied according to electrostatics. At higher pH, they did note a maximum in adsorbed amount at an intermediate potential—consistent with behavior of organics at metal surfaces—and surmised that similar behavior may hold at other pH values, provided a sufficiently large potential range. This nonmonotonic behavior is explained by the hydrophobic effect: at a hydrophobic surface of near-zero charge, the contacting water is highly ordered and thus its displacement is highly favored entropically, whereas at a more charged hydrophobic surface, the contacting water is less ordered and thus more difficult to displace by adsorbing molecules.

Fievet et al.[17] have also used impedance methods to measure protein adsorption onto a carbon electrode. They observed both the adsorbed amount and the initial adsorption rate to exhibit a maximum at an intermediate electrode potential and attributed this as well to a hydrophobic effect.

Bos et al.,[18] using reflectometry, observed a modest influence of an applied electric field on the adsorption of several proteins onto indium tin oxide under most conditions. Most observations followed electrostatic arguments, but an exception was lysozyme at pH=9.9 Fig. 3, where a significant increase in adsorbed amount of this positively

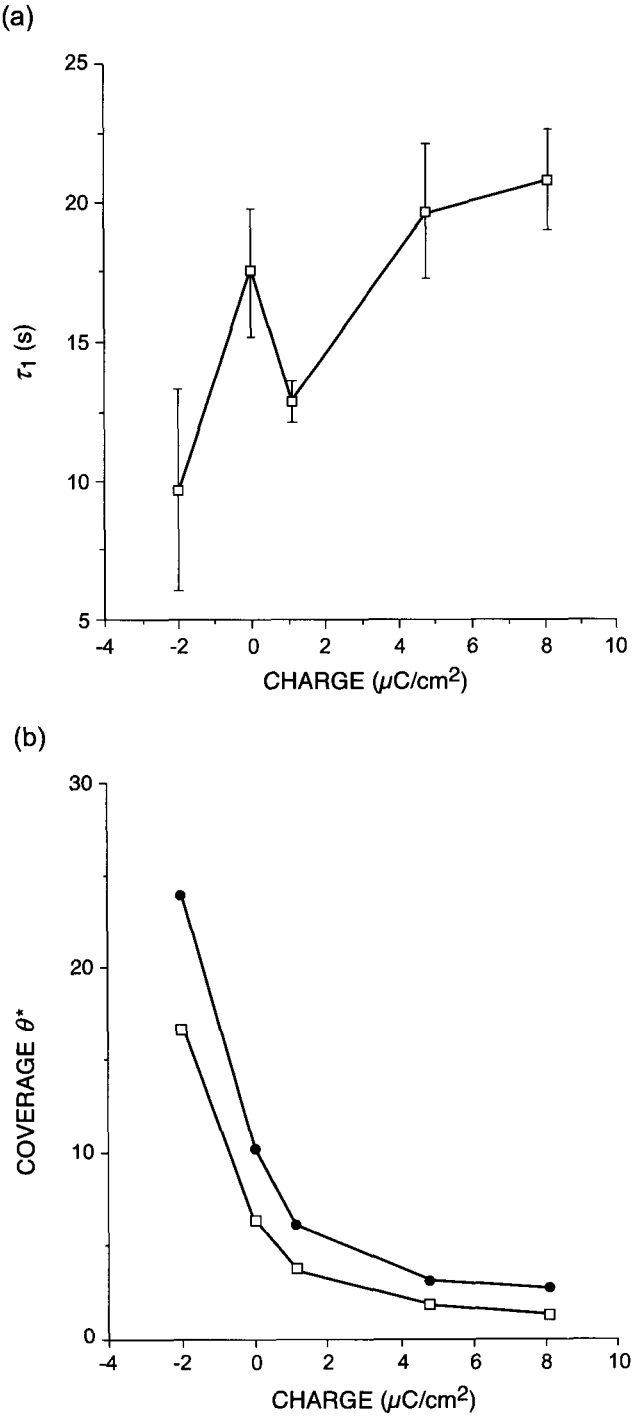


Fig. 1 a) The characteristic time of adsorption and b) the amount adsorbed at saturation in the initially adsorbed state (bottom line) and the conformationally altered state (upper line) vs. electrode potential for fibrinogen adsorbing to a carbon electrode as measured using an impedance method. (From Ref. [15].)

charged protein was observed at ITO potentials exceeding 300 mV (vs. Ag/AgCl). This was explained by a local pH effect: the H^+ concentration decreased with increasing electrode potential, thereby causing the protein to become less positive (or even slightly negative).

Khan and Wernet[19] used a quartz crystal microbalance to investigate the adsorption of glucose oxidase onto an electroconducting polyanion-doped polypyrrole film. Although the study focused on enzymatic activity following adsorption, of particular significance was the demonstration of a kinetic isotherm under an applied potential Fig. 4.

Fraaije et al.[22] have used total internal reflection fluorescence to investigate the voltage dependence of the amount and mean orientation of cytochrome *c* adsorbed onto tin oxide. The fluorescent intensity as a result of absorption of both the p-wave and the s-wave of the incident light was measured, and it was determined that over

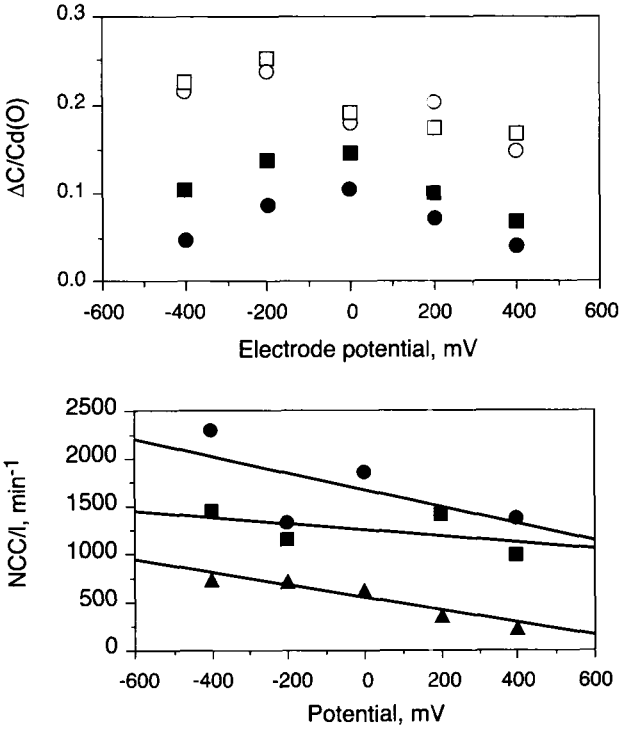


Fig. 2 a) Normalized capacitance change because of adsorption of 0.5 g/L human serum albumin (circles) and 0.1 g/L lysozyme (squares) from acidic (pH=4.0, open) and basic (pH=10.6, solid) solutions as functions of carbon electrode potential as measured by an impedance method. b) Initial rates of normalized capacitance change for 0.50 g/L human serum albumin (circles), 0.05 g/L human plasma fibrinogen, 0.50 g/L superoxide dismutase, and 0.10 g/L lysozyme onto a carbon electrode as a function of electrode potential. The change in interfacial capacitance is a measure of the extent of surface coverage. (From Ref. [16].)

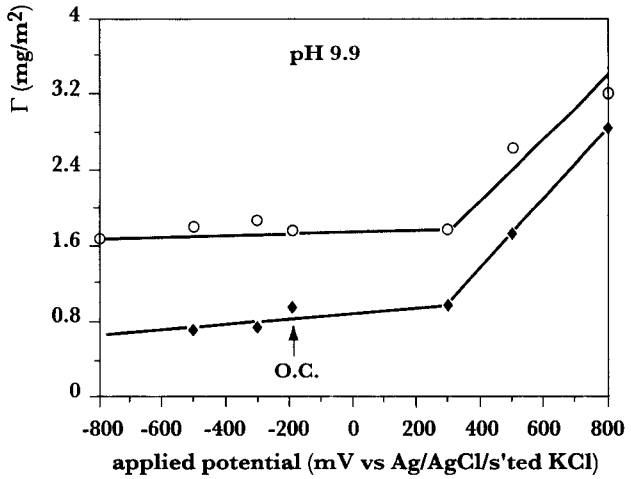


Fig. 3 The amount of lysozyme (open symbols) and ribonuclease (closed symbols) from a solution at pH=9.9, following 15 min of exposure to an indium tin oxide surface, as a function of applied voltage as measured using reflectometry. (From Ref. [18].)

a fairly significant voltage range (from 0 to 1 V vs. Ag/AgCl), little change in adsorbed amount or mean orientation occurs. In contrast, changing the pH or the surface coverage of the adsorbed protein resulted in changes in mean orientation angle 7° and 4°, respectively.

Asanov et al.[23] have used TIRF to measure the adsorption of bovine serum albumin onto tin oxide under an applied electric field. They find little effect on adsorption at anodic potentials (negative surface charge)

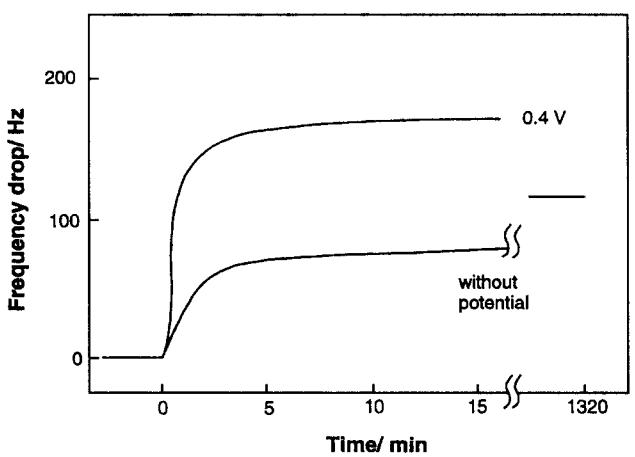


Fig. 4 The amount of glucose oxidase adsorbed to a sulfated poly β hydroxyether-doped polypyrrole film vs. time at an open circuit potential and under an applied potential difference of 0.4 V as measured using a quartz crystal microbalance. (From Ref. [19].)

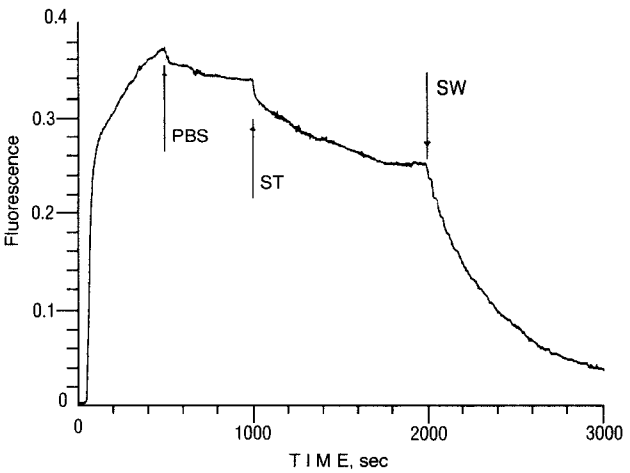


Fig. 5 The fluorescence detected (a measure of the adsorbed amount) from labeled antibiotin adsorbing onto a biotinylated indium tin oxide electrode. At the first arrow, the protein solution is replaced by a pure PBS buffer solution. At the second arrow, the open circuit potential is replaced by a saw tooth polarization from −0.7 to +1.1 V at 0.1 V/sec. At the third arrow, the potential is changed to a square wave polarization from −0.7 to +1.1 V at a time period of 5 sec. (From Ref. [24].)

but a significant increase in adsorption at increasingly cathodic potentials. This behavior again contradicts one's expectations based on electrostatics. However, it was noted that the adsorption behavior correlates strongly with the adsorbing substrate's water contact angle; thus it appears that hydrophobic forces are at play. These authors have also investigated the voltage-induced desorption of biospecifically bound species, with a goal of developing a regenerable biosensing platform.[24] Their findings are impressive: the rate and extent of desorption could be greatly enhanced by the application of an oscillating voltage Fig. 5. In particular, the desorption time constant was decreased from 10^5 sec in the absence of an applied potential to 300 sec in the presence of a square wave potential alternating from −0.7 to +1.1 V.

Brusatori et al.[6,20] have investigated the adsorption of albumin and cytochrome *c* onto ITO using OWLS. The raw data of a typical experiment appear in Fig. 6. Following a baseline measured under flowing solvent (deionized water in this case), application of an electric field (1.0 V over a 1.0-mm distance) between the adsorbing substrate and a platinum counterelectrode (i) yields an increase in the measured effective refractive index, the fundamental output signal of OWLS. This increase is not fully understood, but is likely a result of interfacial double-layer charging, to penetration of small ions into the ITO and/or the underlying silicon titanium oxide films (even in deionized water, carbonate and

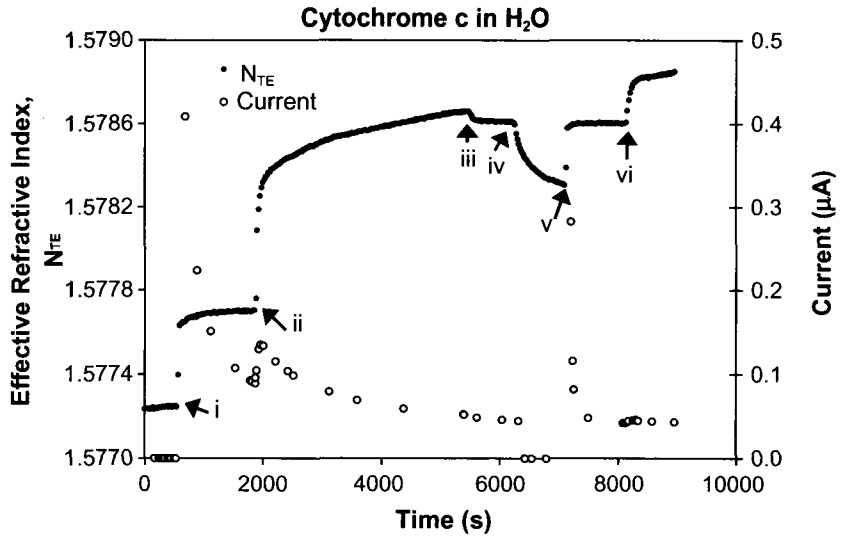


Fig. 6 The effective refractive index (a measure of the adsorbed amount from optical waveguide lightmode spectroscopy) and the current during cytochrome *c* adsorption from water onto indium tin oxide as a function of time. At point i, a voltage difference of 1.0 V is applied. At point ii, the protein solution is introduced. At point iii, the protein solution is replaced by a buffer solution. At point iv, the system is returned to an open circuit potential, and at point v, the 1.0-V voltage difference is reapplied. At point vi, the protein solution is reintroduced. (Taken from Ref. [6].)

other ionic impurities are present), or to mild oxidation of the ITO. The current increases initially but then decreases, as would occur for a resistor and capacitor in series. Upon addition of cytochrome *c* in water (ii), the signal greatly increases because of optical changes at the interface brought about by adsorption. A return to the deionized water (iii) causes only a small signal reduction, indicating that only a small quantity of protein desorbs. When the electric field is removed (iv), a decreased signal results. However, this is not a result of further desorption as a reapplication of the field (v) returns the signal to its previous level. Finally, a reintroduction of the protein (vi) yields an additional signal increase. This additional adsorption is likely a result of an increased amount of area on the surface open for adsorption. The cause of this increased available area is likely aggregation among the adsorbed molecules.

The surface density vs. time for albumin and cytochrome *c*, in a water solvent, adsorbing to ITO under a range of applied voltage appears in Fig. 7. Clearly, adsorption is significantly enhanced under an applied electric field. In fact, at voltages exceeding 1.0 V, no true adsorption plateau is reached. If a rectangular array of adsorbed proteins is assumed, the saturated monolayer density would be 0.46 μg/cm^2. Thus multilayer adsorption is occurring, at least to some extent, in all but the open circuit potential curve of Fig. 7. In Fig. 8, the adsorption rate vs. adsorbed amount is shown. In all curves, an initial increase in adsorption rate is noted, reflecting transport limitations. The subsequent decrease in adsorption rate is a result of surface filling. Thus the initial surface limited kinetics of albumin increase significantly with applied voltage, while those of cytochrome *c* are not affected. At longer times, however, the rate of adsorption for cytochrome *c* also is increased by the presence of the electric field. The transport-limited adsorption rate vs. time (insets) shows that only albumin transport is directly influenced by the field, likely because of its greater overall (negative) charge. While solvent and local pH effects could be influencing these results, the major contributing factor is the preferential adsorption, under an applied field, at patches on the protein surface of complementary charge.

PERSPECTIVES

An applied electric field offers an opportunity to influence the structure and formation kinetics of an adsorbed protein layer. The literature now contains several examples of this.[6,15–20,22–24] However, it is striking the extent to which observed behavior deviates from that expected by basic electrostatics. There is little agreement on the cause of such behavior and therefore on the very nature of the electric field's influence on protein adsorption. Explanations put forth invoke surface-bound counterions, local pH effects, interfacial solvent structure, and protein

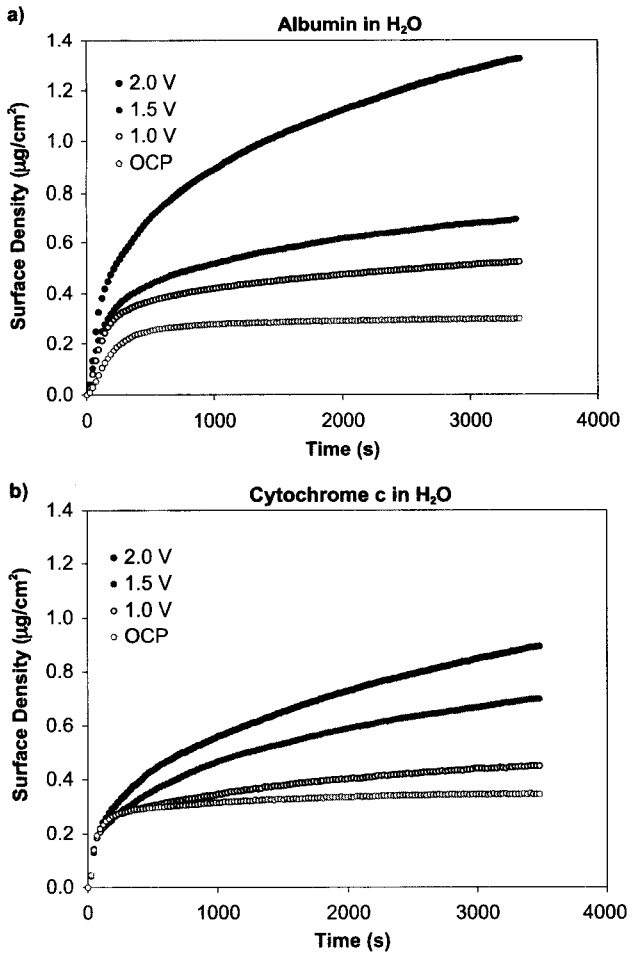


Fig. 7 The amount of a) albumin and b) cytochrome *c* from a water solution adsorbed to an indium tin oxide surface as a function of time for various applied voltages as measured by optical waveguide lightmode spectroscopy. (Taken from Ref. [6].)

charge heterogeneity. Below, we comment upon each of these effects.

Surface-Bound Counterions

Could similarly charged surfaces be attracted to one another given the presence of a third, oppositely charged species? At first glance, this would seem unlikely because counterions would at most completely compensate the charge, leaving two neutral surfaces. The classic Poisson–Boltzmann approach, in which solvated electrolytic species are assumed uncorrelated, predicts only repulsion between like charged objects in an electrolytic solution. However, there are examples where charge overcompensation may occur. A spectacular recent example of this is the layer-by-layer method of growing thin films by exposing a substrate alternatively to solutions of polyanions and polycations.[25] Although not fully understood theoretically, the idea is that charged sites on the polyions exchange nearly completely with smaller ions at the surface for entropic reasons. However, in addition to sites contacting the surface, the polyions generally contain additional charged sites that extend away from the surface, thus resulting in a charge overcompensation. However, this type of charge overcompensation would not be expected from the relatively small ionic species present in most protein adsorption systems.

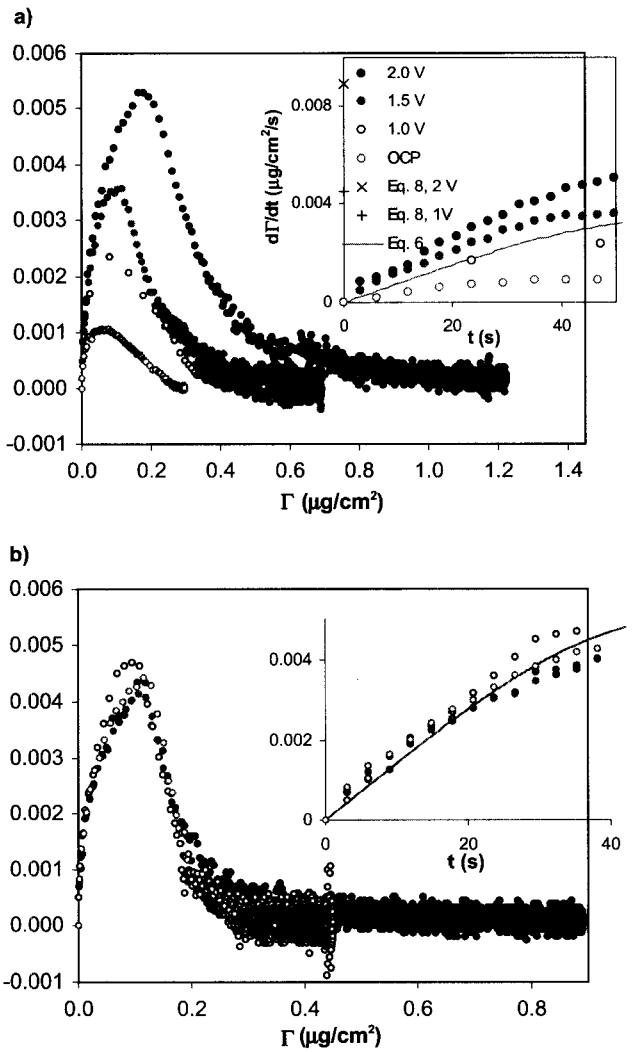


Fig. 8 The rate of adsorption vs. amount adsorbed for a) albumin and b) cytochrome *c* from water onto indium tin oxide as a function of adsorbed amount (and, in the insets, as a function of time) for various applied voltages as measured by optical waveguide lightmode spectroscopy. (Taken from Ref. [6].)

Local pH Effects

Proteins are amphoteric, i.e., possess a number of acid/base sites whose charge is pH-dependent. Because a positively (negatively) charged surface will repel solvated protons (hydroxide ions) and thus raise (lower) the pH near to the surface, the sign of a protein's overall change could change upon approach to the surface. In this case, attraction that is indeed electrostatic in origin may follow. Could this effect explain the antielectrostatic behavior reported in the literature? Evidence to suggest otherwise lies in the following observation: the degree of buffering has been shown to have little effect on these antielectrostatic adsorption trends.[20]

Solvent Interfacial Structure

Solvent molecules (typically water) at or near a solid substrate often experience a hindered rotation relative to those far from the interfacial region. If liberated from the surface via, for example, displacement by an adsorbed protein, a net entropy gain results. An estimation of the rotational contribution to the entropy is given by the statistical mechanical result for an ideal gas rigid rotor, $S_{rot}=Nk \ln q_{rot}$, where N is the number of molecules, k is the Boltzmann constant, and q_{rot} is the rotational partition function. For water at 300 K, $q_{rot}\approx 30$, so the contribution per mole of water liberated from the surface is about 28 J/K mol. An average-sized protein might result in the liberation of 50 water molecules, so the contribution to the overall free energy of adsorption at 300 K from the solvent rotation would be about -420 kJ/mol. Of course, this is an upper bound because the molecules at the surface do possess some rotational freedom. This estimate would most accurately apply to water at a neutral, hydrophobic surface, where the molecules are expected to be very ordered. As the magnitude of surface charge is increased, the water becomes more disordered, i.e., may experience enhanced rotation, and the free energy gain from its liberation is thus diminished. As several of the above-mentioned studies show trends consistent with these thoughts, the role of solvent structure in the overall influence of an applied electric field cannot be discounted.

Protein Charge Heterogeneity

Based on electrostatics, can a positively (negatively) charged protein be attracted to a positively (negatively) charged surface? Within the Poisson–Boltzmann approach, the answer is yes, according to Asthagiri and Lenhoff,[10] so long as both protein and surface contain charge heterogeneity. The idea is that regions of complementary charge may come into close contact, while overall repulsion is sufficiently screened by counterions so that a net attraction occurs. Brusatori et al. invoke this reasoning to explain the increase in adsorption observed for positively charged cytochrome *c* with increasing adsorbent potential. [6]

CONCLUSION

An applied electric field offers the possibility of controlling the structure of an adsorbed protein layer. This represents a powerful method of producing nanostructured coatings for biosensing and biomaterials applications. Only in the past few years have techniques become available to measure the adsorbed layer properties in situ during formation. While it remains to fully resolve a number of puzzling observations—and while quantitative prediction of the effects of the electric field, such as counterion binding, local pH enhancement, solvent structure, and charge heterogeneity, remains sparse—the prospects are strong for a general understanding, in the not-to-distant future, of the influence of an electric field on protein adsorption, and the engineering of protein coatings to near-exact specifications using voltage-based methods is certainly a realistic and laudable goal.

ACKNOWLEDGMENTS

The author thanks the National Institutes of Health, under Grant # R01-EB00258, for financial support.

REFERENCES

1. Haynes, C.A.; Norde, W. Globular proteins at solid/liquid interfaces. Colloids Surf., B Biointerfaces **1994**, *2*, 517–566.
2. Malmsten, M. Formation of adsorbed protein layers. J. Colloid Interface Sci. **1998**, *207* (2), 186–199.
3. *Biopolymers at Interfaces*; Malmsten, M., Ed.; Surfactant Science Series, Marcel Dekker: New York, 1998; Vol. 75.
4. Van Tassel, P.R. Biomolecules at Interfaces. In *Encyclopedia of Polymer Science and Technology*, 3rd Ed.; Wiley Interscience: New York, 2003; 285–305.
5. Calonder, C.; Van Tassel, P.R. Kinetic regimes of protein adsorption. Langmuir **2001**, *17* (14), 4392–4395.
6. Brusatori, M.A.; Tie, Y.; Van Tassel, P.R. Protein adsorption kinetics under an applied electric field: An optical waveguide lightmode spectroscopy study. Langmuir **2003**, *19* (12), 5089–5097.

7. Tie, Y.; Calonder, C.; Van Tassel, P.R. Protein adsorption: Kinetics and history dependence. J. Colloid Interface Sci. **2003**. *268* (1), 1–11.
8. Roth, C.M.; Lenhoff, A.M. Electrostatic and Van-Der-Waals contributions to protein adsorption—Comparison of theory and experiment. Langmuir **1995**, *11* (9), 3500–3509.
9. Roth, C.M.; Lenhoff, A.M. Electrostatic and Vanderwaals contributions to protein adsorption—Computation of equilibrium-constants. Langmuir **1993**, *9* (4), 962–972.
10. Asthagiri, D.; Lenhoff, A.M. Influence of structural details in modeling electrostatically driven protein adsorption. Langmuir **1997**, *13* (25), 6761–6768.
11. Oberholzer, M.R.; Lenhoff, A.M. Protein adsorption isotherms through colloidal energetics. Langmuir **1999**, *15* (11), 3905–3914.
12. Ravichandran, S.; Talbot, J. Mobility of adsorbed proteins: A Brownian dynamics study. Biophys. J. **2000**, *78* (1), 110–120.
13. Carignano, M.A.; Szleifer, I. Adsorption of model charged proteins on charged surfaces with grafted polymers. Mol. Phys. **2002**, *100* (18), 2993–3003.
14. Fang, F.; Szleifer, I. Competitive adsorption in model charged protein mixtures: Equilibrium isotherms and kinetics behavior. J. Chem. Phys. **2003**, *119* (2), 1053–1065.
15. Bernabeu, P.; Caprani, A. Influence of surface-charge on adsorption of fibrinogen and or albumin on a rotating-disk electrode of platinum and carbon. Biomaterials **1990**, *11* (4), 258–264.
16. Feng, L.; Andrade, J.D. Protein adsorption on low-temperature isotropic carbon. 2. Effects of surface-charge of solids. J. Colloid Interface Sci. **1994**, *166* (2), 419–426.
17. Fievet, P.; Mullet, M.; Reggiani, J.C.; Pagetti, J. Influence of surface charge on adsorption of a hydrophobic peptide onto a carbon surface by capacitance measurements. Colloids Surf., A Physicochem. Eng. Asp. **1998**, *144* (1–3), 35–42.
18. Bos, M.A.; Shervani, Z.; Anusiem, A.C.I.; Giesbers, M.; Norde, W.; Kleijn, J.M. Influence of the electric potential of the interface on the adsorption of proteins. Colloids Surf., B Biointerfaces **1994**, *3* (1–2), 91–100.
19. Khan, G.F.; Wernet, W. Adsorption of proteins on electro-conductive polymer films. Thin Solid Films **1997**, *300* (1–2), 265–271.
20. Brusatori, M.A.; Van Tassel, P.R. Biosensing under an applied voltage using optical waveguide lightmode spectroscopy. Biosens. Bioelectron. **2003**, *18* (10), 1269–1277.
21. Bearinger, J.P.; Voros, J.; Hubbell, J.A.; Textor, M. Electrochemical optical waveguide lightmode spectroscopy (EC-OWLS): A pilot study using evanescent-field optical sensing under voltage control to monitor polycationic polymer adsorption onto indium tin oxide (ITO)-coated waveguide chips. Biotechnol. Bioeng. **2003**, *82* (4), 465–473.
22. Fraaije, J.; Kleijn, J.M.; Vandergraaf, M.; Dijt, J.C. Orientation of adsorbed cytochrome-C as a function of the electrical potential of the interface studied by total internal-reflection fluorescence. Biophys. J. **1990**, *57* (5), 965–975.

Protein Adsorption Studied by Atomic Force Microscopy

David T. Kim
Harvey W. Blanch
Clayton J. Radke
University of California, Berkeley, California, U.S.A.

INTRODUCTION

Nonspecific protein adsorption onto solid surfaces occurs in many important applications. These include drug storage, where adsorption of a therapeutic protein onto the container walls may denature and render the drug inactive, as well as change its solution concentration, posing a major problem when exact dosing is required.[1] Protein adsorption can contribute to blood clotting and heart disease;[2] the biocompatibility of materials used in medical implants, food processing, and chromatographic media depends on the interaction of proteins with the surface.[3–8] Newly developed biosensors and immunoassay methods depend on the immobilization of proteins.[9] Fouling of contact lenses is due, in part, to the nonspecific adsorption of protein on the lens surface.[10] In all of these applications, a fundamental understanding of protein adsorption is required to control the interaction between the protein and the solid.

This article discusses the role of atomic force microscopy (AFM) in elucidating protein adsorption mechanisms onto solid surfaces. After a brief review of other experimental techniques, we outline the advantages and disadvantages of AFM for in situ study of protein adsorption and summarize findings for AFM available to date. We conclude with a brief discussion of the protein adsorption process.

OVERVIEW

Measurement Techniques

A wide range of techniques provides information on protein adsorption onto solid surfaces. Vibrational sum frequency spectroscopy (VSFS)[a] is a surface-specific nonlinear optical technique in which a sum frequency signal is obtained from an ordering of dipoles at surfaces. Moieties that are well aligned at the surface can be observed by VSFS. Using this technique, information on protein organization in addition to water structure and organization can be obtained. In particular, Kim et al.[11] used this technique to study the organization of lysozyme at the water/quartz interface at various pH values.

The interaction force between two solid surfaces can be measured as a function of distance by the surface force apparatus (SFA). Claesson et al.[12] employed SFA to study a wide range of proteins (globular, unordered, fibrous) and determined protein conformation, orientation, and the operative forces. Small compact globular and soft globular proteins could be distinguished by measuring their compressibility. Blomberg et al.[13] used SFA to study the adsorption of lysozyme on mica as a function of protein concentration, determining the protein's adsorption orientation, ability to form multilayers, and adsorption irreversibility.

Conformation information on adsorbed protein is also available from several spectroscopic methods. Circular dichroism spectroscopy (CD), Fourier transform infrared spectroscopy (FTIR), and Raman spectroscopy all probe protein secondary structure. These techniques have been useful in demonstrating how proteins alter their structure upon adsorption. Proteins have been shown to have greater structural perturbation on hydrophobic surfaces, compared to hydrophilic surfaces using CD,[14,15] FTIR,[7] and Raman spectroscopy.[8,16]

A significant amount of literature on protein adsorption is concerned with the kinetics and total mass of protein adsorbed onto a solid surface. The quartz crystal microbalance (QCM) measures the changes in resonance frequency and dissipation factor of an oscillating quartz crystal and can provide information on the adsorbed mass and temporal variations in surface viscoelastic properties. Otzen et al.[17] employed QCM to study the adsorption of protein S6 onto a methyl-terminated quartz surface and found that the adsorption kinetics of protein S6 depends on the equilibrium fraction of denatured protein in the bulk, rather than on the kinetics of bulk denaturation. Upon comparison with optical techniques, such as ellipsometry and optical waveguide lightmode spectroscopy (OWLS), Hook et al.[18] showed that QCM reports higher adsorbed mass, this being attributed to water bound to the adsorbed protein.

[a]A glossary of abbreviations appears at the end of this article.

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120014115

Total internal reflection fluorescence (TIRF) measures the fluorescence emission of fluorescent-tagged protein triggered by an evanescent wave arising from a totally internally reflected light beam inside an adjacent optically clear adsorption substrate. This method permits adsorption isotherms and adsorption kinetics to be measured. It has also been used to show that proteins reorient and/or undergo conformational change at the surface, depending on the surface concentration,[19] the rate of adsorption to the surface,[19] and the hydrophobicity of the surface.[20,21] These factors all affect the final surface loading.[22]

Other optical methods exist to determine adsorption kinetics and isotherms. Ellipsometry gauges the change in polarization of light that is reflected from the surface of a sample. This technique requires a reflecting surface. In OWLS, light is totally internally reflected within an optical waveguide, setting up an evanescent wave in the nearby solution. The electric and transverse magnetic modes of the evanescent wave are then measured. In surface plasmon resonance (SPR), a laser beam undergoes total internal reflection at the adsorption glass / metal interface. Coupling between the photon and a surface plasmon causes excitation of surface plasmon at the interface and a reduction in intensity of the internally reflected light. Scanning angle reflectometry (SAR) is another optical method that is used to determine protein adsorption isotherms and kinetics. All these techniques measure the adsorption thickness and the refractive index of the layer above the adsorption interface, which can then be used to calculate adsorbed mass. Surface plasmon resonance[23] and ellipsometry[24] show how adsorption packing and kinetics vary for proteins of various dimensions and conformations. Ellipsometry has shown that more protein adsorbs onto hydrophobic surfaces compared to hydrophilic surfaces.[25] Ellipsometry also has been used to show that more protein is irreversibly adsorbed on hydrophobic surfaces[26] and that the degree of irreversibility is time dependent.[25,26] Time-dependent spreading of protein at the interface has been studied using OWLS[27] and ellipsometry.[25] Multilayer formation for lysozyme has been observed using ellipsometry.[24,28] Optical techniques have been used to study how ionic strength affects protein adsorption area,[29] adsorption isotherms,[28,30] kinetics,[28] and irreversibility.[28]

Atomic Force Microscopy Techniques

All of the above experimental techniques provide information on length scales much larger than the dimensions of a single protein molecule. Hence they cannot provide data on the spatial distribution of proteins on a surface. Conversely, atomic force microscopy (AFM) enables direct observation of real-space topographic surfaces on a nanometer scale and can, in principle, detect the spatial distribution of adsorbed protein. Atomic force microscopy can image surfaces in solution, so that proteins remain in an aqueous environment. By imaging in situ, adsorption dynamics are observable. By combining the real-time and real-space resolution of AFM, one obtains information on the spatial and temporal distribution of adsorbed protein. The literature on AFM studies of proteins is vast. However, most studies are primarily concerned with imaging large, isolated single-protein molecules[31–33] and not on adsorption patterns and mechanisms. We review studies on the fundamentals of the protein adsorption process. Table 1 contains a listing of AFM research that we mention in this article, along with the particular properties studied.

Access to lateral spatial information on protein adsorption enables supramolecular organization on the surface to be observed in real time. This includes the ability to directly observe clustering and multilayer formation. Height differences across the surface provide information on the adsorbed protein conformation and orientation and their changes in time. Nonuniform surfaces can be imaged, allowing comparison between protein adsorption on differing surface chemistry and on patterned interfaces. Atomic force microscopy provides information on the interaction forces between the protein-covered surface and the scanning probe and mechanical properties using the force-distance-curve feature.

ATOMIC FORCE MICROSCOPY

Features

Originally designed to measure the topography of surfaces, AFM has the potential to obtain atomic-scale spatial resolution. Although specific designs differ, AFM generally operates by raster scanning a fine, cantilevered tip over a solid surface while measuring tip deflection with a laser. As the tip encounters features on the surface, the tip deflects; the feedback mechanism moves the sample closer or farther from the cantilever-tip assembly in response. This height movement, in addition to the raster scanning movement, is controlled to angstrom precision by a piezoelectric crystal. It is the height of the sample surface (rather than the tip deflection) that is reported as the height of features on the surface.

Two main methods for imaging topographical features exist. In "contact mode," the tip is kept in constant contact with the surface while maintaining a constant force (or deflection). The instrument feedback loop operates by raising or lowering the sample surface to maintain a constant tip force. This is the original method

Table 1 Atomic force microscopy studies highlighted in this article

First author	Protein	Surface	Real time	Multilayer	Protein structure and orientation	Nonuniform surfaces	Force curves	Comments	Reference
Chen, X. et al.	BSA, BFG	PS					×	Change in adhesion as protein film forms, breakthrough force to surface increases with time	[67]
Coen, M.C. et al.	Protein A	Silicon		×	×			Multilayer formation over time, subsequent layers become less structurally perturbed	[55]
Cullen, D.C. et al.	IgG and GOx	HOPG	×		×			Comparison of adsorption mechanisms for two structurally different proteins	[56]
Galli, G. et al.	Protein A and F-actin	Si and Ti, nanostructures				×		Effects of nanostructures on protein adsorption	[62]
Kim, D.T. et al.	Lysozyme	Mica	×	×	×		×	Complete monolayer formation monitoring, multilayer formation, tip contamination issues	[45]
Mueller, H. et al.	Cytochrome *c*, MBP	Lipid bilayers	×					Adsorption behavior on mixed acidic, basic, and zwitterionic lipid bilayers	[54]
Radmacher, M. et al.	Lysozyme	Mica					×	Classic force curve work on lysozyme, analysis of adhesion, estimation of viscosity and Young's modulus	[66]
Raghavachari, M. et al.	VWF	Mica and OTS/glass			×			Hydrophobicity determines adsorbed protein structure	[59]
Sheller, N.B. et al.	HAS	CH_3 SAMs			×			Protein coverage at surface determines adsorbed protein structure	[48]
Ta, T.C. et al.	BFG	HOPG and mica	×		×			Hydrophobicity determines adsorbed protein structure	[37]
Ta, T.C. et al.	BFG	CH_3 and COOH SAMs	×		×	×		Hydrophobicity determines adsorbed protein structure	[44]
Takahara, A. et al.	BSA	OTS/FOETS mixed monolayer	×			×		Adsorption behavior on mixed monolayer at different pH conditions	[51–53]

See Abbreviation Glossary.

of operation for the AFM; it has the advantage of high resolution. Continuous improvements in preparation techniques and imaging conditions make it possible to image proteins with lateral resolution of ~0.5 nm and vertical resolution of ~0.1 nm in this mode.[34,35] "Tapping-mode" imaging was developed to reduce possible damage to the sample that results from the tip "dragging" along the surface. This is especially relevant for "soft" biomolecules such as proteins. Tips can damage protein structure and sweep away loosely adsorbed biomolecules. In tapping-mode imaging, the cantilevered tip oscillates at resonance frequency and only intermittently "taps" the surface at the bottom of its oscillation. In this imaging mode, the feedback loop maintains a constant oscillation amplitude, rather than the tip force. Using tapping mode, proteins that are loosely adsorbed can be imaged without being swept away by the scanning tip.[31] The lateral resolution suffers somewhat in this mode, but is still sufficient to monitor protein adsorption.

A major advantage of AFM is that imaging can be performed in liquid solution. Thus protein can be imaged in aqueous solution. With a fluid flow cell, the liquid environment can be changed while the sample remains on the surface. Thus protein solution, surfactant solution, or different buffer solutions can be introduced. Imaging adsorbed proteins after drying the surface is a common application. Although sample preparation and imaging in open air tends to be easier and faster, we caution that the drying process may distort protein structure and pattern formation. In most, but all of the images presented in "Atomic Force Microscopy and Protein Adsorption," the tapping mode in aqueous environment is employed.

Other imaging variations for the AFM exist. One mode, known as lateral- (or friction-) force mode can "image" a force map by detecting differences in chemistry and mechanical properties at the surface. This is possible by operating the AFM in contact mode and measuring the "twisting" of the cantilever as the tip moves over different regions of the sample. Twisting of the cantilever is a result of the tip adhering to the surface as it is scanned; the amount of twisting is a function of the adhesion between the tip and surface. A resulting friction map distinguishes parts of the surface with different chemistry and/or mechanical properties.

In addition to imaging, AFM is able to measure the interaction normal force between the tip and the surface. This occurs by measuring the tip deflection as the surface approaches the probe tip, comes into contact, slightly pushes the tip, then retracts and separates from the tip. Information is garnered on the physical and mechanical properties of surfaces and the interaction forces between the tip and surface. When the interaction between the surface and tip material is not of direct interest, tips can be coated with proteins or monolayers that expose a specific chemical functionality to probe-specific interactions. This is sometimes referred to as chemical force microscopy.[36] If force curves are taken at regular distances over a surface, a so-called force-volume (actually force-area) map can be created. In a force-volume map. the tip deflection from a force curve at a defined tip–sample distance is mapped into a two-dimensional array. This discloses the tip interactions with the surface over various regions of the surface. The force-volume map can be compared to the topographical surface images and physical features on the surface can be correlated to the force interactions.

Imaging Artifacts

Care should be taken when interrogating AFM images. Interpreting images without cognizance of possible artifacts can lead to misinterpretation. Because AFM is an intrusive process, the probe tip may mechanically induce protein adsorption patterns on the surface, possibly contributing to the overall adsorption process. Many researchers have reported the "sweeping" action of the tip. Ta et al.[37] and McMaster et al.[38] both reported streaky images attributable to tip mechanical influence. Denis et al.[39] found that elongated structures were bent toward the left or right depending on the scanning direction of the tip. Johnson et al.[40] attributed bare patches devoid of adsorbed molecules to the tip sweeping away protein molecules. Immunoglobulin G (IgG)[41] and immunoglobulin M (IgM)[42] are known to aggregate on mica as a result of tip mechanical influence.

Contamination of the AFM tip by unwanted particles in solution or particles garnered from the surface is a common problem and is difficult to avoid. Depending on the scanning size, the effects on imaging may be significant. Tip contamination alters the tip–surface interaction force and may have a significant effect on the height information obtained.[43] In Ta and McDermott's[44] studies with friction-force microscopy, the authors observed that tip contamination probably occurred and affected the magnitude of the tip friction. Fortunately, their studies were directed at friction contrast, rather than absolute values. In our work with in situ lysozyme adsorption onto mica, tip contamination distorted the apparent size of the protein adsorbed onto the surface, compounding the problem of tip broadening.[45]

With tip broadening, the AFM image is enlarged as a result of a convolution of the imaged particle and the imaging tip.[46,47] Tip broadening is significant when the size of the imaged particle is of the same length scale of the tip point, which is the case for proteins. Enlargement of the image causes overestimation of the amount of protein adsorbed onto the surface, making quantitative assessment of adsorption amounts problematic. Sheller et al.[48] have shown that a grain size analysis of the unmodified image of albumin adsorption yielded a value of 10–15% coverage. However, after taking into account tip

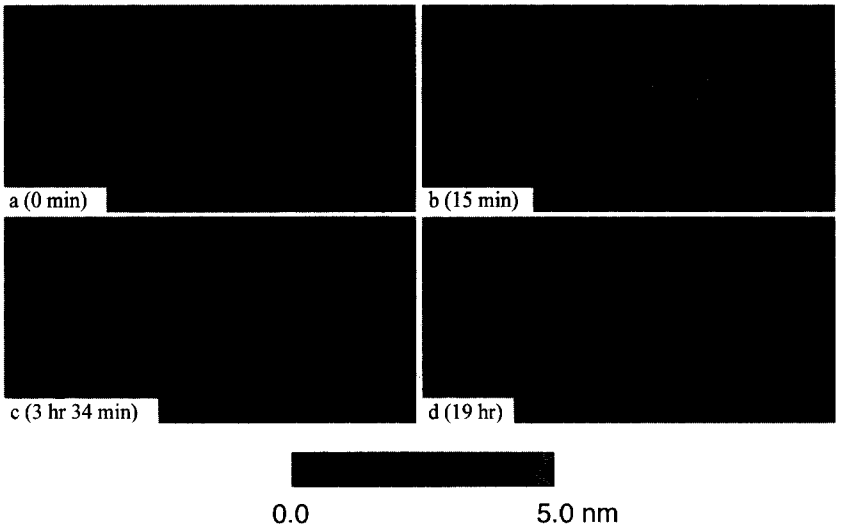


Fig. 1 Effect of tip contamination on AFM imaging. Atomic force microscopy images of adsorption of HEWL on mica in 10 mM acetate buffer, pH 4.0, under stopped-flow conditions. Initial bulk concentration of HEWL is 2 μg/mL. Each image was taken at the exact same location. A large prominent white feature (i.e., debris) marks the same location on the mica surface. (a) Bare mica before exposure to protein solution. (b) Lysozyme adsorption after 15 min, (c) 3 hr 34 min, and (d) 19 hr. Note that in this last scan, the particles below the large feature are much smaller than those above the feature and those scanned at earlier times. All images are 250 nm × 500 nm. Reprinted from Ref. [45] with permission from The American Chemical Society.

broadening, the coverage was reduced to 3–4%. In Ta et al.'s[37] studies with fibrinogen adsorption, it is difficult to identify early adsorption features as single molecules or clusters. Thus these researchers considered the adsorbed particles to be clusters of protein. In the work of Kim et al.[45] with lysozyme and mica, tip contamination contributed to the tip-broadening problem. This can be seen in Fig. 1. We initially interpreted the observed enlargement of the protein features on the surface as a result of cluster growth over time. However, this was disproved when after 19 hr (Fig. 1d), smaller clusters were again observed, presumably after the tip was inadvertently "cleaned" of debris. There exist methods to characterize the shape of the tip so that the amount of broadening can be estimated.[49,50] However, the amount of contamination can change at any time during imaging or exposure to protein solution, as documented in Kim et al.[45]

Finally, we remark that AFM is limited to surfaces that are smooth on the length scale of the protein molecules or aggregates. Accordingly, solid surfaces studied to date include mica, graphite, silicon (i.e., actually a SiO_2 overgrowth), selected polymers, and coated monolayers on these surfaces.

ATOMIC FORCE MICROSCOPY AND PROTEIN ADSORPTION

One of the main advantages of AFM over other techniques is its ability to gather information on the supramolecular organization of adsorbed protein. Atomic force microscopy, unlike optical techniques such as OWLS, SPR, TIRFS, or ellipsometry, has superb lateral spatial resolution. Further, unlike electron microscopy, sample preparation is relatively easy. Protein adsorption events can be observed in situ in aqueous environments and in real time.

Supramolecular Organization

Real-time imaging

Takahara et al.[51–53] observed a real-time increase of bovine serum albumin (BSA) coverage on surfaces covered with *n*-octadecyltrichlorosilane (OTS) islands surrounded by 2-perfluorooctyltrichlorosilane (FOETS) at both pH 4.7 and 7.5. Complete monolayer coverage can be obtained with sufficient protein in solution. Mueller et al.[54] imaged myelin basic protein (MBP) adsorption onto negatively charged, acidic lipid bilayers supported on mica and observed that protein elevations on the surface grew laterally and eventually fused with neighboring protein elevations to form a complete monolayer. As reported in Fig. 2, hen egg white lysozyme (HEWL) formed clusters on mica that increased in number until a complete monolayer was formed.[45] Coen et al.[55] saw similar adsorption patterns for protein A adsorbing onto silicon.

Because of its excellent lateral resolution, AFM can distinguish between different adsorption patterns. Cullen and Lowe[56] examined how protein and surface structures

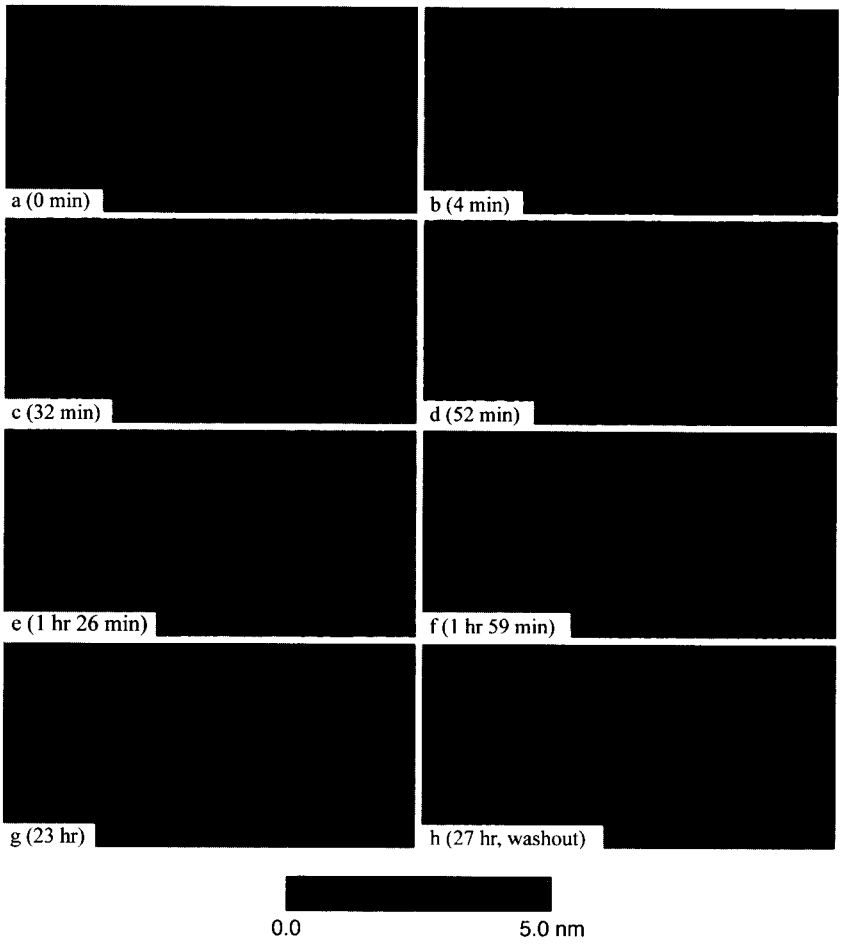


Fig. 2 Lysozyme adsorption onto mica. Atomic force microscopy images of adsorption of HEWL on mica in 10 mM acetate buffer, pH 4.0, under stopped-flow conditions. The initial bulk concentration of HEWL is 5 μg/mL. Each image is 500 nm × 100 nm. The light areas are clusters of protein molecules and the dark areas represent the bare mica surface. (a) Bare mica surface before contact with the protein solution. (b) Lysozyme adsorption after 4 min, (c) 32 min, (d) 52 min, (e) 1 hr 26 min, (f) 1 hr 59 min, and (g) 23 hr. (h) Washout with buffer after 27 hr. (Reprinted from Ref. [45] with permission from The American Chemical Society.)

dictate adsorption patterns by adsorbing IgG and glucose oxidase (GOx) onto highly oriented pyrolytic graphite (HOPG), as shown in Figs. 3 and 4. Within minutes, patches of IgG containing 300–500 molecules were observed across the HOPG plane (Fig. 3B). Over time, more patches at random locations appeared on the surface until a complete monolayer formed (Fig. 3D). Conversely, in Fig. 4, GOx nucleated initially at HOPG defects. Subsequent adsorption occurred at locations adjacent to protein already adsorbed, forming "sheets" of protein and leaving regions of bare HOPG. In both cases, the protein appeared to aggregate on the surface, which is likely a result of hydrophobic residues exposed during adsorption conformation change leading to favorable intraprotein interaction at the surface. The differences in adsorption patterns arise from physiochemical differences between the two proteins, which we will discuss later.

Conversely, Ta et al.[37,44] demonstrated how differences in the hydrophobicity of the solid surface affect protein adsorption patterns. Fig. 5 displays consecutive AFM images of bovine fibrinogen (BFG) film growth on HOPG. At 75 sec (Fig. 5C), a network of branched strands can be observed extending across the HOPG plane with no apparent preference for step defects. After 125 sec (Fig. 5E), more networks of branched strands are apparent. A complete monolayer is formed after 225 sec (Fig. 5G). Fig. 5H shows a friction image of the complete monolayer, indicating that the frictional forces between the tip and protein film are low and constant, meaning that the protein structure is robust and crystalline-like.

On mica, as depicted in Fig. 6, BFG homogeneously adsorbed across the surface until a full monolayer was formed. This adsorption process is very similar to that of lysozyme on mica shown in Fig. 2. The authors also

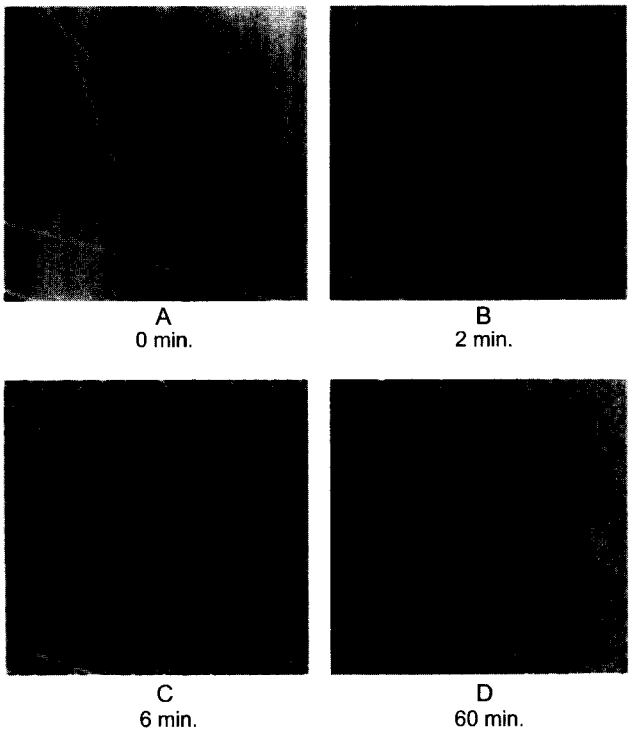


Fig. 3 Immunoglobulin G adsorption onto HOPG. Time-resolved AFM images of the adsorption of IgG from a phosphate-buffered saline (PBS) solution with a bulk concentration of 50 μg/mL of protein to a freshly cleaved HOPG surface. Image size: (A–C) 5.3 μm × 5.3 μm, (D) 5 μm × 5 μm. All images were obtained with an applied repulsive force of ~1 nN and presented as left-shaded topographic images. (Reprinted from Ref. [56] with permission from Elsevier.)

compared BFG adsorption onto octadecanethiol (ODT; CH_3-terminated) and mercaptohexadecanoic acid (MHA; COOH-terminated) self-assembled monolayers (SAMs) supported on gold-covered glass.[44] The hydrophobic, CH_3-terminated surface showed similar adsorption behavior to that of the HOPG surface, with strand-like surface features gradually filling in the surface to form a smooth, well-packed, uniform monolayer. Friction-force images showed that the friction properties across the surface did not significantly change during the adsorption process. However, the hydrophilic, COOH-terminated surface showed different adsorption behavior to that of the mica surface. Instead of homogeneously filling the surface as in Fig. 6, on the -COOH surface, circular, island-like domains 4-nm in height appeared on the surface and grew laterally, ultimately forming a complete, well-packed layer. Friction-force images in this case showed friction contrast between the adsorbed protein and bare surface. The discrepancy between the adsorption patterns on the hydrophilic mica and the COOH-terminated surface may be because of the specific chemical differences at the interface such as polarity and charge density and distribution.

Multilayer formation

With AFM, one can visually determine the existence of multilayers and when and how they occur. Kim et al.[45] showed that lysozyme formed a second layer of adsorption (Fig. 2g), only after the first monolayer was complete (Fig. 2f). In addition, formation of the second layer was significantly slower than the first. Coen et al.[55] showed that protein A, when adsorbed on silicon, also formed multilayers as pictured in Figs. 7 and 8. Fig. 7a shows an incomplete monolayer of protein A, with the dark spots indicating locations of bare silicon. The existence of a single incomplete monolayer in Fig. 7 was confirmed using X-ray photoelectron spectroscopy (XPS). As with lysozyme adsorption on mica, subsequent protein layers seen in Fig. 8a did not form until the

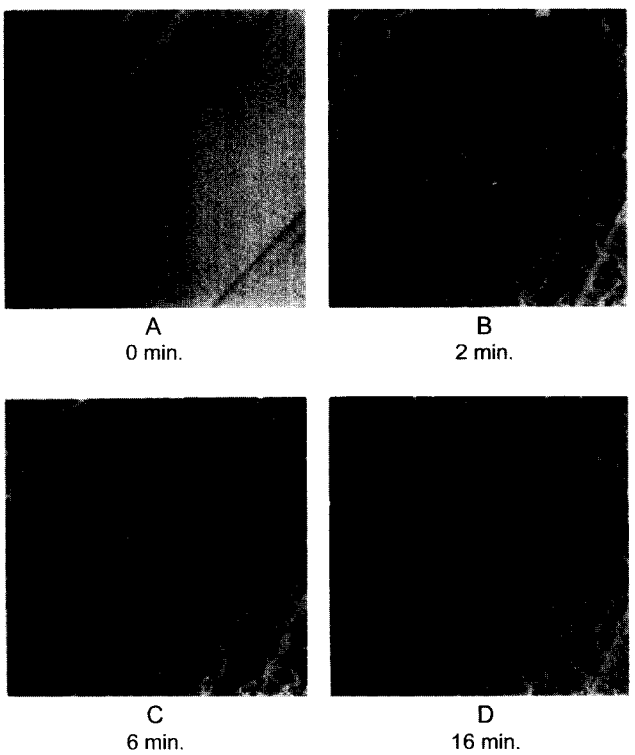


Fig. 4 Glucose oxidase adsorption onto HOPG. Time-resolved AFM images of the adsorption of GOx from a phosphate-buffered saline (PBS) solution with a concentration of 50 μg/mL of protein to a freshly cleaved HOPG surface. All images are 5 μm × 5 μm in area, imaged with an applied repulsive force of ~1 nm and presented as left-shaded topographic images. (Reprinted from Ref. [56] with permission from Elsevier.)

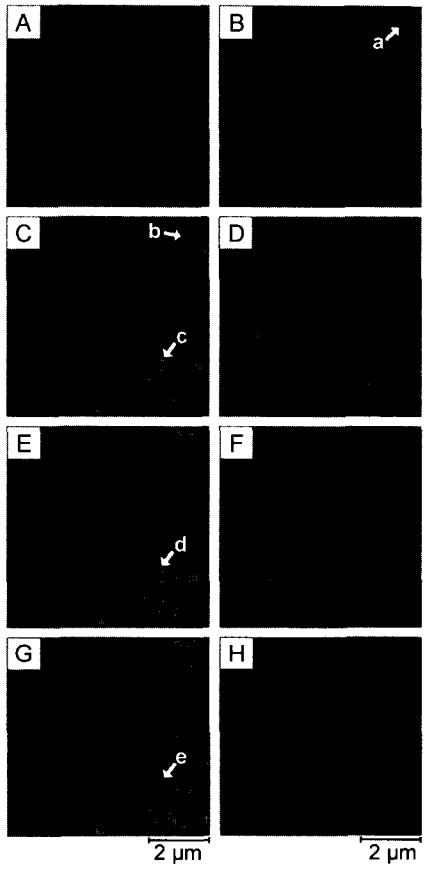


Fig. 5 Bovine fibrinogen adsorption onto HOPG. 6.6 μm × 6.6 μm consecutive AFM images of BFG film growth on HOPG monitored in real time with continuous flow. Parts A and B are topographic images (z-scale = 10 nm). Parts C, E, and G are sequential topographic images (z-scale = 10 nm) while parts D, F, and H are the corresponding lateral force images (z-scale = 0.2 V). For timing purposes, the center of the scan in part B is defined as $t=0$, as this is the point initial adsorption is observed. C and D were captured 75 sec after initial adsorption. E and F correspond to 125 sec, and G and H correspond to 225 sec. Arrow a highlights a BFG strand that elongates to a branching point shown in arrow b. Arrow c demarks the growth of a single aggregate later fully formed and highlighted by arrow e. Arrow d indicates a defect site on the HOPG plane. (Reprinted from Ref. [37] with permission from The American Chemical Society.)

previous monolayer was complete. The cases of lysozyme on mica and protein A on silicon differ in their electrostatic conditions. At pH 4.0, mica is negatively charged while the lysozyme is positively charged. For the protein A / silicon system, at pH values between 5 and 6, the protein is slightly negatively charged and the surface is also negatively charged. Despite differences in charge conditions, both systems displayed multilayer formation. The role of electrostatics in protein adsorption is covered later.

Adsorbed Protein Structure and Orientation

Using AFM, it is possible to ascertain adsorbed protein structure and orientation. Although height information (i.e., adsorbate layer thickness) is available from other experimental techniques, AFM has the advantage of directly imaging differences in terrain height with nanometer lateral resolution. Optical techniques average over hundreds of nanometers.

Several researchers have shown with AFM that proteins exhibit structural changes and lateral spreading at the surface. Bovine fibrinogen adsorbed onto ODT (CH_3-terminated) SAMs displays an adsorbed AFM height of 1–1.5 nm that is in contrast to the adsorbed height of 4 nm on MHA (COOH-terminated) SAMs.[44] The smallest dimension of the BFG molecule obtained from X-ray crystallography is 4–6 nm. The lower height on the CH_3-terminated SAMs is indicative of more extensive conformational spreading on that surface, which is a result of its more hydrophobic nature. Glucose oxidase was also shown to have a lower adsorbed height of 5–8 Å on HOPG[56] compared to the dimensions of the native protein ($7.0 \times 5.5 \times 8.0$ nm).[57] However, not all proteins show extensive lateral spreading on hydrophobic surfaces. The thickness of IgG films on HOPG was shown to be 3 nm, compared to the dimensions of the protein given by X-ray crystallography ($14.2 \times 8.5 \times 3.8$ nm).[58] In fact, using AFM, Raghavachari et al.[59] have shown an apparently opposite behavior. Fig. 9 illustrates the hydrophilic mica surface after 1 hr of exposure to von Willebrand factor (VWF) followed by protein-free phosphate buffer saline (PBS) washout. The VWF on this surface adopted a spread-out conformation. On the hydrophobic OTS surface, shown in Fig. 10, VWF maintained a "ball of yarn" configuration.

Different amounts of spreading were observed for subsequent layers in the multilayer adsorption of protein A on silicon.[55] The first monolayer exhibited a height of approximately 1 nm. This is shown in Fig. 7b by the section line of the silicon surface after 15 sec of adsorption. The size of native protein A is roughly 3 nm.[55] Thus a significant amount of lateral spreading (and likely unfolding) occurred in the first layer of adsorbed protein, assignable to direct contact with the silicon surface. The height of the second monolayer was considerably larger at 2–3 nm (Fig. 8b), indicating less structural perturbation. The height on the third monolayer was approximately 3 nm (not shown.) Thus Coen et al.[55] concluded that the protein in the third layer was not conformationally altered. Each new monolayer adsorbs onto a different surface. Therefore it is not surprising that the adsorbed protein structure differs in each layer. However, given very long time exposure to the surface, protein in the multilayers may continue to rearrange, slowly unfolding and aggregating into the sublayer.

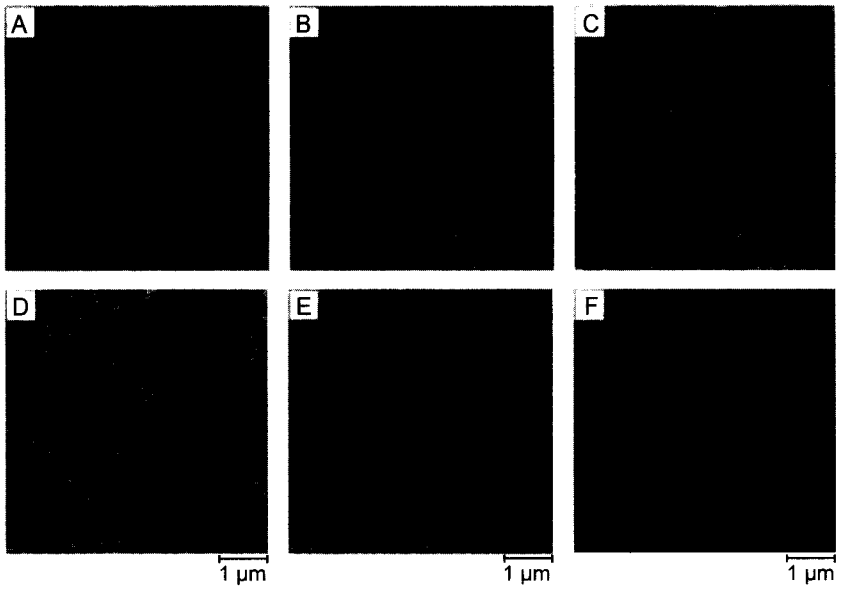


Fig. 6 Bovine fibrinogen adsorption onto mica. Atomic force microscopy images, 5 μm × 5 μm, of BFG film growth on mica monitored in real time with continuous flow. All images are topographic with a z scale = 10 nm except for part A, where the z scale = 5 nm. (A) Upward scan $t = 0$ sec is defined at halfway through the scan. (B) Successive downward scan, with an elapsed time of 75 sec. (C) Successive upward scan with an elapsed time of 125 sec. (D) Successive downward scan with an elapsed time of 175 sec. (E) Successive upward scan with an elapsed time of 225 sec. (F) Upward scan with an elapsed time of 325 sec. (Reprinted from Ref. [37] with permission from The American Chemical Society.)

In many cases, protein molecules were observed to adsorb in a side-on orientation on solid surfaces. Kim et al.[45] showed that the height of adsorbed lysozyme on mica was approximately 2.5 nm while the native protein dimensions from X-ray crystallography are 3 × 3 × 4.5 nm.[60] The height of adsorbed C hordein on mica is 2–3 nm[38] while the dimensions of the protein are 1.9 nm in width by 28.2 nm in length.[61] By adsorbing in a side-on fashion, the protein maximizes contact with the surface.

There are cases where protein does not adsorb in a side-on fashion. Sheller et al.[48] examined human serum albumin (HSA) adsorption onto hexadecyltrichlorosilane (HTS; CH_3-terminated) SAMs and found that the adsorbed orientation depends on surface concentration. For low submonolayer coverage, as shown in Fig. 11c,d, the height profile of protein adsorbed onto the surface measured by AFM was 3.6–3.7 nm. Protein films of a full monolayer, as shown in Fig. 11a,b, had a thickness of 16–17 Å as measured by X-ray reflectivity. Sheller et al. suggested that for higher surface coverage of protein, there is a significant net removal of weakly adsorbed protein from the surface by adjacent, strongly adsorbed protein spreading on the surface. This implies that the stronger adsorbed protein undergoes structural changes to a greater extent giving more contacts with the surface. When the surface coverage is low, the exchange process is not significant, and proteins adsorb in any configuration. Thus both strongly and weakly bound protein adhere to the surface with varying degrees of conformation change. Sheller et al. do not provide experimental evidence that weakly bound protein remains on the surface for long exposure times.

Nonuniform Surfaces

The exceptional lateral resolution of AFM allows study of nonuniform and patterned surfaces and of the differences in protein adsorption as a direct result of surface nonuniformities. Patterned surfaces presenting both hydrophobic and hydrophilic regions highlight the differences in the degree of hydrophobicity on protein adsorption.[44] Fig. 12 shows a topographical (A) and friction image (B) of a protein-free surface, patterned with ODT and MHA SAMs. These images illustrate that the two SAMs are topographically rather similar, but the friction-force image establishes that the tip experiences different adhesive forces on the two SAMs. Fig. 13 presents friction images of the patterned surface before (Fig. 13A) and during exposure to bovine fibrinogen (BFG) solution (Fig. 13B–D). Image D was collected after 22 min of exposure to the protein solution and again shows that the tip experiences different adhesive forces now on the fibrinogen layers of each type of SAM. This result indicates that the protein absorbs on both surface regions, but in a different conformation or orientation on the hydrophobic (CH_3-terminated) and hydrophilic (COOH-terminated) surfaces.

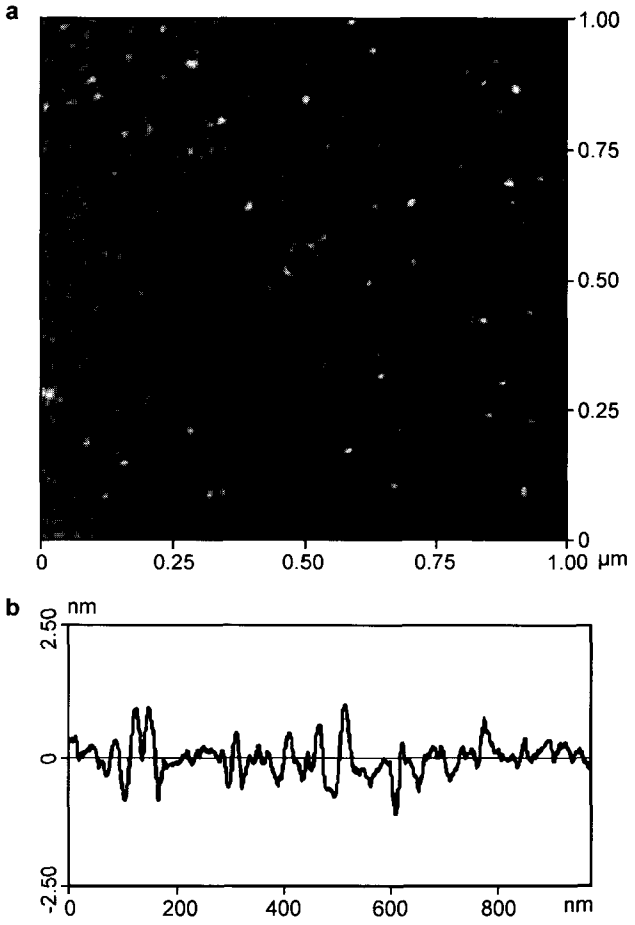


Fig. 7 Protein A adsorption onto silicon. (a) Atomic force microscopy picture of the silicon surface after a 15-sec adsorption of protein A from a 0.1 mg/mL solution. The first monolayer of protein A is almost complete. Bare mica is evidenced by the black spots in the image. The vertical scale is 5 nm. (b) Line section of the AFM image displaying the height of features. (Reprinted from Ref. [55] with permission from Elsevier.)

Patterned surfaces also highlight differences in adsorption behaviors on different surfaces with varying electrostatic conditions.[51–53] Fig. 14a shows a mixed monolayer consisting of OTS surrounded by FOETS supported on a silicon substrate. When exposed to a BSA solution at pH 7.5, only the FOETS regions demonstrated significant adsorption, as illustrated in Fig. 14b. However, at pH 4.7, the isoelectric point of BSA, the protein adsorbed to both phases, as demonstrated in Fig. 14c. Although BSA adsorbed onto both OTS and FOETS surfaces when uncharged, it has an apparently higher affinity for FOETS, which is, apparently, more hydrophobic than OTS. At pH 7.5, the protein carries a net negative charge and readily adsorbed on FOETS. Adsorption onto the OTS phase was possibly prevented by electrostatic repulsion from already adsorbed negatively charged protein on the FOETS phase.[51–53]

The effects of the physical differences in terrain on protein adsorption can also be elucidated by AFM. Galli et al.[62] examined the effect of nanostructures created by local anodic oxidation on protein adsorption. They created elevated lines 30–40 nm in width and 1–3 nm in height on silicon (Si) and titanium (Ti) surfaces (actually SiO_2 and TiO_2 overgrowths, respectively). Fig. 15 shows an AFM image of the 1-nm high lines on Si. Fig. 16 shows that the lined nanostructures (highlighted with arrows) made no difference to protein A (a globular protein with a size of 3 nm) adsorption in terms of adsorbed amount. In contrast,

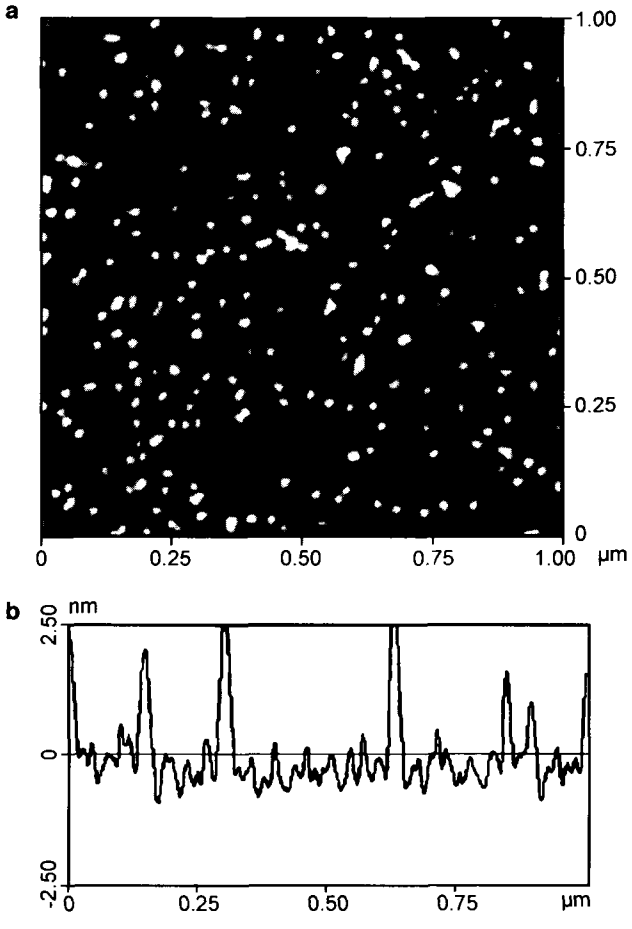


Fig. 8 Protein A multilayer adsorption onto silicon. (a) Atomic force microscopy picture of the silicon surface after an 8-min adsorption of protein A from a 0.1 mg/mL solution. A second layer of adsorbed protein is evident as the brightest spots. The vertical scale is 5 nm. (b) Line section of the AFM image displaying the height of features. (Reprinted from Ref. [55] with permission from Elsevier.)

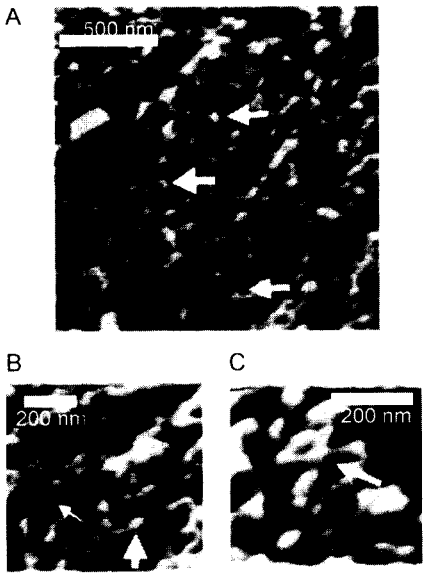


Fig. 9 von Willebrand factor adsorption onto mica. A is an AFM image (2 μm × 2 μm scan) of VWF multimers adsorbed on hydrophilic mica and imaged under PBS. Most of the VWF molecules are extended and are not ball of yarn structures. Extended VWF multimer chains overlap with themselves and with neighboring chains (arrows). Intramolecular structural features in the VWF multimer chain are discernible. Individual chains of multimers are not clearly distinguishable because of the high surface density of VWF. Mean lateral dimensions of the VWF multimers are 640±161 nm × 303±113 nm. B and C are expanded AFM images of VWF on mica, showing intramolecular structural features of individual chain VWF multimers. VWF chains are composed of globular domains (B, large arrow) joined by thin rod-like interconnecting sections (B, small arrow). C contains an example of a crossover (see arrow) of the VWF multimer chain within the same VWF chain. (Reprinted from Ref. [59] with permission from Elsevier.)

F-actin, a filamentous protein with a diameter between 6.5 and 8.2 nm, had a tendency to adsorb along the 1-nm-high nanostructured lines, whereas no preferential orientation occurred on the flat surfaces. Fig. 17 shows F-actin adsorbed on the Si surface containing nanostructures 1 nm in height. Fig. 18 shows F-actin adsorption on Ti with 1- and 3-nm nanostructures and with no nanostructures. Thus on Si and Ti, F-actin had a tendency to adsorb along 1-nm-high nanostructured lines, whereas no preferential orientation occurred on the planar surfaces. Galli et al. speculated that the reason for preferential orientation is that the contact area between the surface and protein is increased along the lines. However, on Ti, 3–4-nm-high lined nanostructures garnered very low F-actin adsorption. This height corresponds to half the diameter of F-actin. These authors cite similar surface height-dependent adsorption for a variety of cells.[63–65] It is possible that F-actin is not sufficiently labile to adsorb across these higher nanostructures and maintain enough contacts with the surface to adsorb.

Force Interactions

A major strength of AFM is the probing of interactions between the protein-coated surface and the tip. Force interactions can be mapped across different regions of the surface. Radmacher et al.[66] illustrated the utility of force–distance curves for protein adsorption. Fig. 19a and b present force curves on bare mica and on aggregates of lysozyme on mica in 10 mM KH_2PO_4 (pH 6.4), respectively. Force curves report the cantilever deflection (which can readily be translated into force) as a function of the sample height as the tip approaches, and then retracts from the surface. The large (~25 nm) negative deflection of the retract curve on bare mica indicates that the tip is highly attracted to the surface, whereas the smaller (~8 nm) negative deflection over lysozyme indicates a smaller attraction. An identical result was found by Kim et al.[45] The deviation in linearity on the contact region of the force–distance curve for lysozyme is a result of the elastic deformation of the protein. By fitting to a Hertzian model, Radmacher et al.[66] estimated the Young's modulus for adsorbed lysozyme on mica to be 0.5±0.2 MPa. In addition, the time for liftoff, the region of the retract force curve when the tip loses contact with

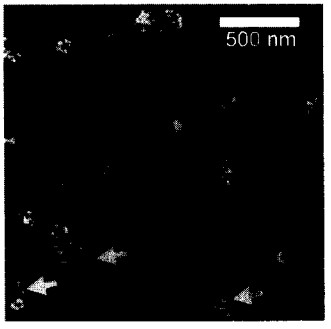


Fig. 10 von Willebrand factor adsorption onto OTS. Atomic force microscopy image (2 μm × 2 μm scan) of VWF multimers adsorbed on hydrophobic OTS and imaged under PBS. Most VWF multimers display the characteristic compact ball of yarn structures observed by electron microscopy. Each VWF multimer is closely packed with intramolecular overlap and crossover of chains. Intramolecular structural features of the VWF multimer chain are resolved. However, the compact arrangement of the chain makes it difficult to discern the structural features belonging to neighboring repeat units. In some multimer chains, short sections are not as compact and appear extended (arrows). Completely extended chains are rare. None are seen in this image area. The average lateral dimensions of the VWF multimers are 256±74 nm × 152±62 nm. (Reprinted from Ref. [59] with permission from Elsevier.)

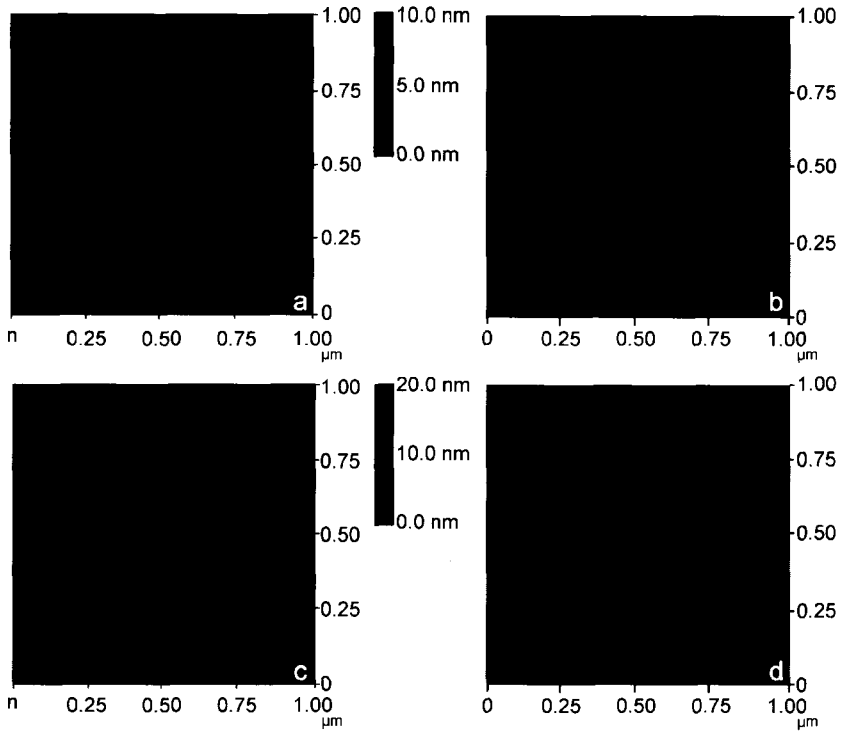


Fig. 11 Human serum albumin adsorption onto HTS. Atomic force microscopy images of substrates covered with HTS SAMs with HSA layer adsorbed at different bulk protein concentrations: (a) 0.5 mg/mL; (b) 0.1 mg/mL; (c) 0.05 mg/mL; (d) 0.01 mg/mL. Note that the z range in images a and b is one-half that in images c and d. Images a and b represent full monolayer coverage while images c and d show incomplete monolayer coverage. (Reprinted from Ref. [48] with permission from The American Chemical Society.)

the surface, is longer for lysozyme compared to that of the bare mica surface. This is related to the viscosity of lysozyme on the surface, estimated to be 800 ± 400 Pa sec.[66]

Clearly, AFM can detect mechanical properties of protein films. Given a high-enough applied force, the tip can break through protein films. Thus by examining the force curves, one can uncover protein assembly at the surface. On myelin basic protein (MBP) clusters adsorbed at both acidic (30 wt.% cholesterol, 56% egg phosphatidylcholine, 14% dioleoylphosphatidylserine) and basic (30 wt.% cholesterol, 56% egg phosphatidylcholine, 14% dioleoyl trimethyl-ammonium salt) lipid bilayers supported on mica, Mueller et al.[54] observed a "jump-in" on the initial approach curve that disappeared for subsequent force curves at the same location. The adhesion force between the tip and the cluster was approximately 0.5 nN as measured by tip deflection upon retraction. Fig. 20b shows a typical set of consecutive force curves on an acidic lipid bilayer. Force–distance curves consistently displayed a characteristic "jump-in" upon approach at 5 nN and a negligible adhesion force. The "jump-in" indicates the moment when the tip breaks through the lipid bilayer and protein layer (if present) to the solid surface. Fig. 20c also gives four consecutive force–distance curves but now for a MBP cluster adsorbed on the acidic lipid bilayer. The disappearance of the jump after the first force curve in Fig. 20c indicates that the tip locally destroyed the protein structure and that the original structure did not recover. The tip may irreversibly compress and break apart the protein aggregate structure, leaving behind a permanently damaged protein–lipid bilayer complex and expose the bare surface. On the bare lipid layer, as shown in Fig. 20b,

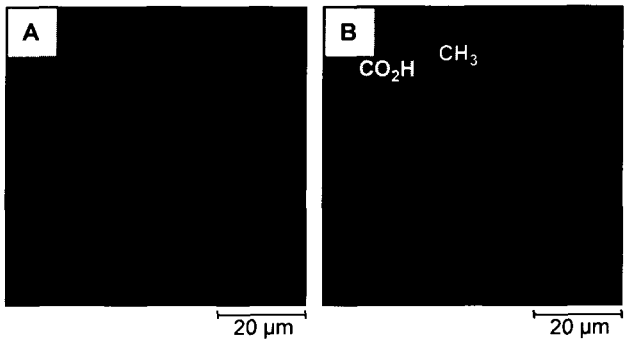


Fig. 12 Octadecanethiol and mercaptohexadecanoic acid nanopatterns on gold. Images, 70 μm × 70 μm, collected in ambient air, of a patterned monolayer prepared by microcontact printing ODT and back-filling with MHA. (A) Topography (z scale 5 nm). (B) Friction image (z scale 0.5 V). (Reprinted from Ref. [44] with permission from The American Chemical Society.)

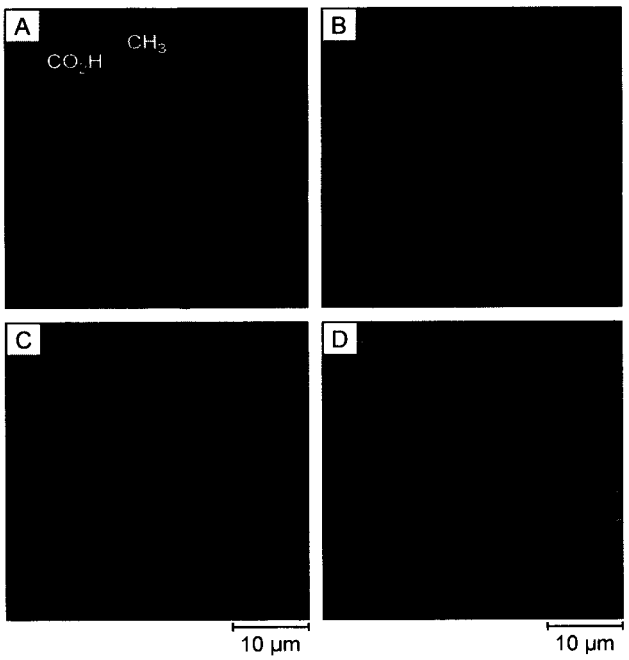


Fig. 13 Bovine fibrinogen adsorption onto nanopatterned ODT/MHA SAMs. Friction images, 40 μm × 40 μm, depicting the real-time adsorption of BFG to a patterned monolayer. All images were collected under flowing solution. (A) Initial surface under flowing PBS. Methyl groups show higher contrast in images A (z scale 0.7 V), B (z scale 0.7 V), and C (z scale 0.4 V) during the adsorption of BFG. (D) Completed BFG film (z scale 0.3 V). Image D was collected 22 min after initial addition of BFG to the fluid cell. (Reprinted from Ref. [44] with permission from The American Chemical Society.)

the jump-in consistently occurred and the recovery after tip breakthrough occurred fast, probably because of the high mobility of the lipid molecules. On zwitterionic bilayers (not shown), protein clusters showed inconsistent behavior. This is not surprising because the protein aggregates on zwitterionic bilayers were irregular in height and two-dimensional morphology.[54]

Chen et al.[67] found long-term structural changes in BSA films adsorbed on polystyrene (PS) surfaces. Albumin-coated probes exhibited high adhesion forces on bare PS, but only small adhesion forces on protein-exposed surfaces. When high probe forces were exerted on the BSA films, high adhesion forces were observed, indicating that the albumin-coated probe was able to penetrate the protein layer to the bare PS surface. These authors found that the contact-force limit increases with protein adsorption time. Hence after adsorption, protein continues to undergo structural changes that increase the mechanical strength (or integrity) of the film.

Force interactions between the tip and the surface at a fixed distance away from the surface can allow presentation of a two-dimensional array or a force-volume image. Fig. 21a shows a force-volume map and the corresponding topograph (Fig. 21b) of a BSA film adsorbed onto PS.[67] The brighter shades on the force-volume map indicate higher adhesion energy. These correspond to the low area of the topograph corresponding to bare PS. Kim et al.[45] employed force-volume imaging for a lysozyme submonolayer on mica to distinguish between areas that corresponded to bare mica and to lysozyme.

Force–distance curves can be used to detect differences in adsorbed protein structure. Denis et al.[39] studied collagen adsorption on CH_3- and OH-terminated SAMs using

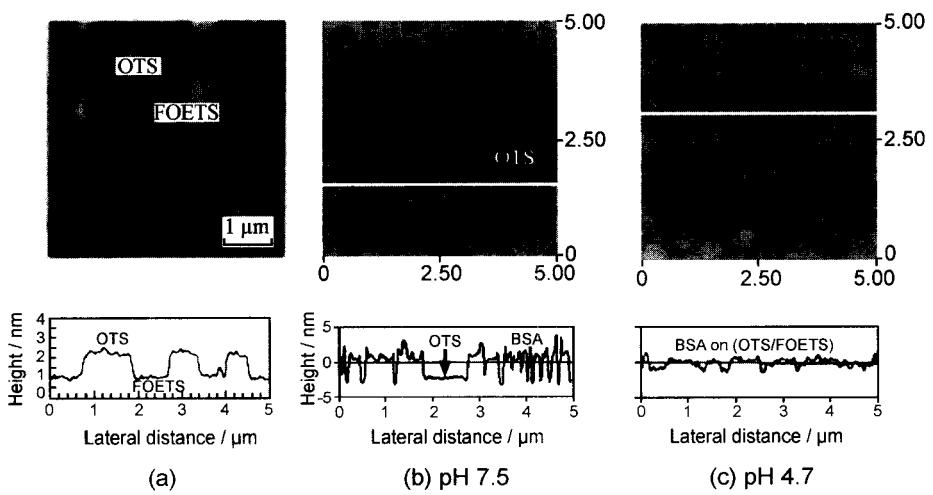


Fig. 14 Bovine serum albumin adsorption on OTS/FOETS supported monolayers. Atomic force microscopy image of the mixed (OTS/FOETS) monolayer (a) before exposure to BSA, (b) after exposure to a 0.1 g/L BSA solution in PBS (0.01M) at pH 7.5 and (c) pH 4.7 for 10 min. Line sections are shown below each AFM image. (Reprinted from Ref. [53] with permission from Elsevier.)

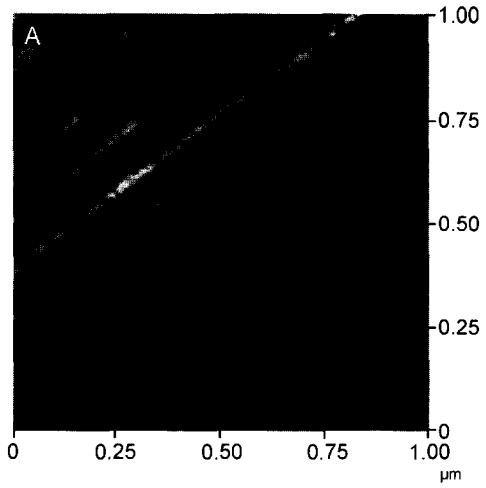


Fig. 15 Nanostructured Si. Atomic force microscopy image of Si surface with 1-nm high lines. The vertical scale is 5 nm. (Reprinted from Ref. [62] with permission from Elsevier.)

force–distance curves. Approach interaction curves for protein adsorbed onto CH_3-terminated SAMs exhibited significant curvature, whereas the approach curves for protein adsorbed onto-OH-terminated SAMs demonstrated no curvature. This result indicates that collagen adsorbed on the hydrophobic CH_3-terminated SAMs was "softer" than that adsorbed on the hydrophilic OH-terminated SAMs. Differences in protein adsorption behavior on hydrophobic and hydrophilic surfaces are addressed below.

PROTEIN ADSORPTION MECHANISMS

Electrostatic Effects

Electrostatic interactions between the protein and surface in addition to those between proteins can play an important role in the adsorption onto solid surfaces. Many adsorption processes are driven in part by electrostatic attraction between the surface and protein. For example, lysozyme at pH 4.0 is positively charged, whereas mica is negatively charged; protein adsorption occurs quite readily as shown in Fig. 2.[31,45,66] Cytochrome *c* at pH 7.4 is positively charged and significantly adsorbs on negatively charged bilayers, minimally on neutral bilayers, and not at all on positively charged bilayers.[54] In addition, cytochrome *c* adsorption was reversible under high salt conditions (5 M NaCl) whereas the protein appeared to be irreversibly adsorbed at lower salt conditions (150 mM NaCl). In some cases, protein can still adsorb onto surfaces even when they are similarly charged. MBP at pH 7.4 is positively charged and adsorbs onto a positively charged bilayer.[54] Patel et al.[68] have shown that negatively charged catalase at pH 8.0 adsorbs onto negatively charged SAMs. Likewise, HSA can adsorb to both positively and negatively charged polymer surfaces.[30] There are a several reasons why this can occur. First, a protein of a particular net charge still presents oppositely charged residues on its structure and those residues can favorably interact with the charged solid surface. Second, as discussed later, the driving force for adsorption usually involves hydrophobic forces. These can overcome a net electrostatic repulsion with the surface.

High-ionic-strength solutions shield electrostatic repulsions between charged proteins resulting in increased protein adsorption. Johnson et al.[40] illustrated this with ferritin adsorption onto trimethoxysilylpropyldiethylenetriamine (DETA) coated glass. However, adsorption amount does not monotonically increase with ionic strength because of the competition between surface–protein interactions and surface protein–protein interactions.[40] Protein–protein repulsion can also be minimized by adjusting the solution pH to the isoelectric point (p*I*) of the protein, where the protein has a zero net charge, but may have a large dipole moment. Several researchers have reported that adsorption is increased or maximized at the p*I*.[40,51–53,69,70] Kajiyama et al. observed that BSA adsorbed on both OTS and FOETS phases only at the p*I* (Fig. 14c). But when BSA carried a net charge, it only adsorbed only onto the FOETS phase (Fig. 14b) as a result of the electrostatic repulsion between adsorbed protein and protein in solution.[51–53] Johnson et al.[40] reported that at the p*I*, ferritin adsorption approached the random sequential adsorption (RSA) jamming limit. This increased

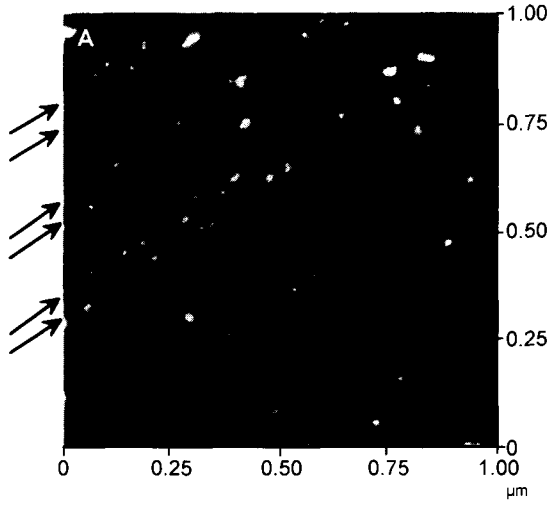


Fig. 16 Protein A adsorption on nanostructured Si. Atomic force microscopy image of protein A adsorption on the Si nanostructures shown in Fig. 15. The vertical scale is 5 nm. Arrows highlight placement of lines. (Reprinted from Ref. [62] with permission from Elsevier.)

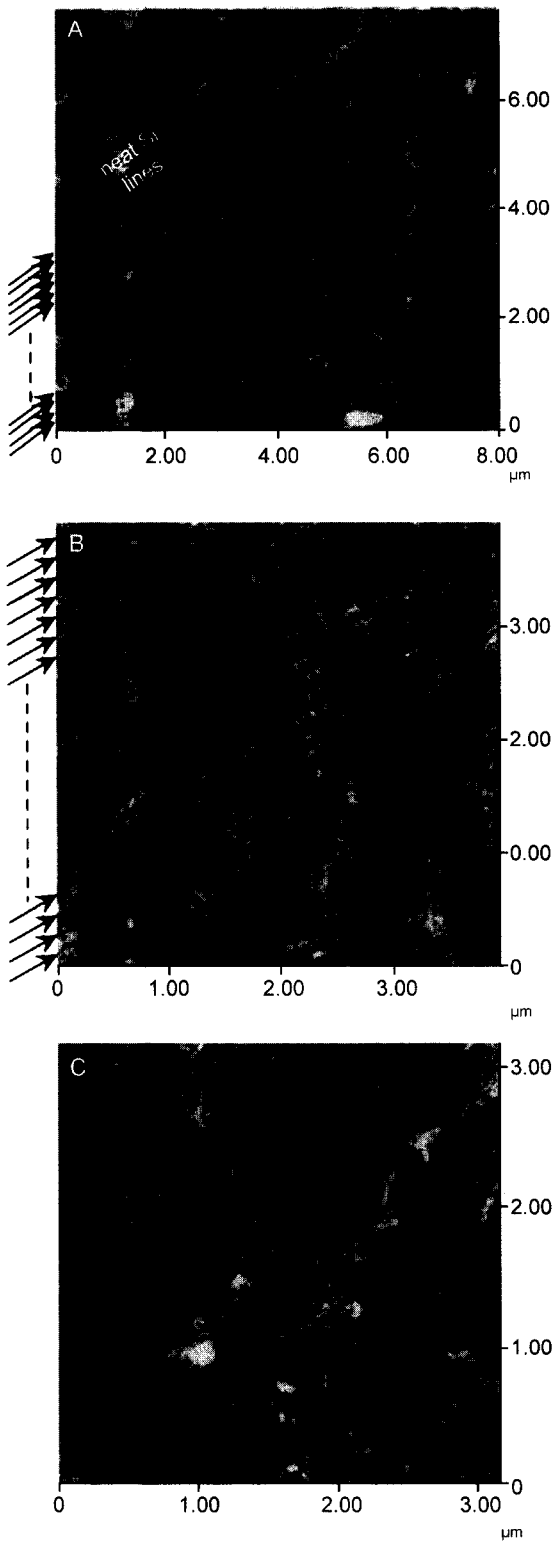


Fig. 17 F-actin adsorption on Si nanostructures. Adsorption of F-actin on surface displaying both Si nanostructures and planar Si as shown in Fig. 15. The vertical scales are 10 nm. (A) General view of the adsorbed F-actin. (B) Enlargement of the nanostructure area. (C) Enlargement of the planar Si area. Arrows highlight placement of lines. (Reprinted from Ref. [62] with permission from Elsevier.)

adsorption could be attributed to the fact that at the p*I*, intramolecular charge repulsions are minimized and the protein adopts its most compact shape.[69] Thus the protein requires less space on the surface to adsorb as compared to a more elongated structure that exists when the protein carries a net charge. However, this reasoning must be tempered by the observation that denatured and random-coil-like proteins adsorb more quickly[17] most likely because the surface unfolding kinetic step is hastened.

Kinetics of Adsorption

General trends have been observed in the kinetics of protein adsorption. Almeida et al.[70] divided the process into four stages, assuming a continuous supply of protein. The first stage is a mass-transfer-limited stage where proteins rapidly adsorb onto the surface. Feng et al.[2]

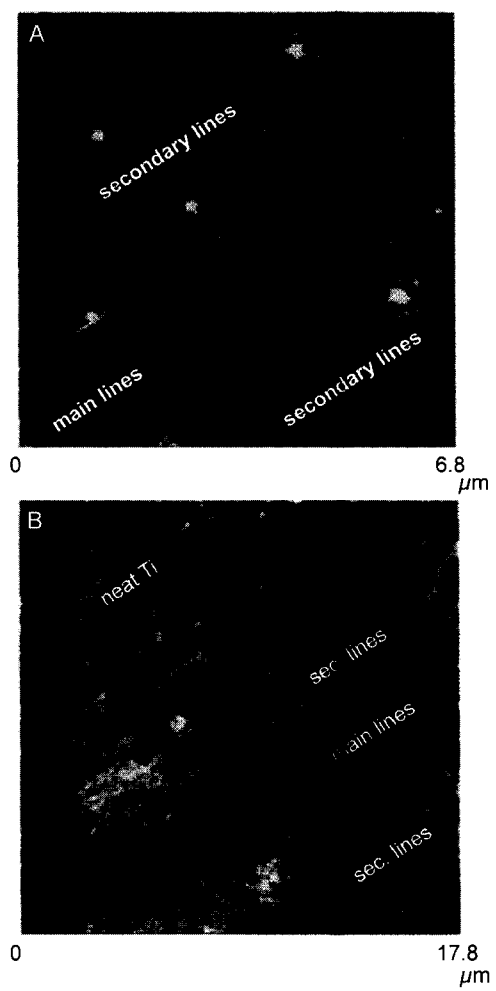


Fig. 18 F-actin adsorption on Ti nanostructures. (A) Nanostructures created on Ti. Main lines are 3–4 nm high. Secondary lines are 1 nm high. (B) F-actin adsorbed on nanostructured Ti. The vertical scales are 8 nm. (Reprinted from Ref. [62] with permission from Elsevier.)

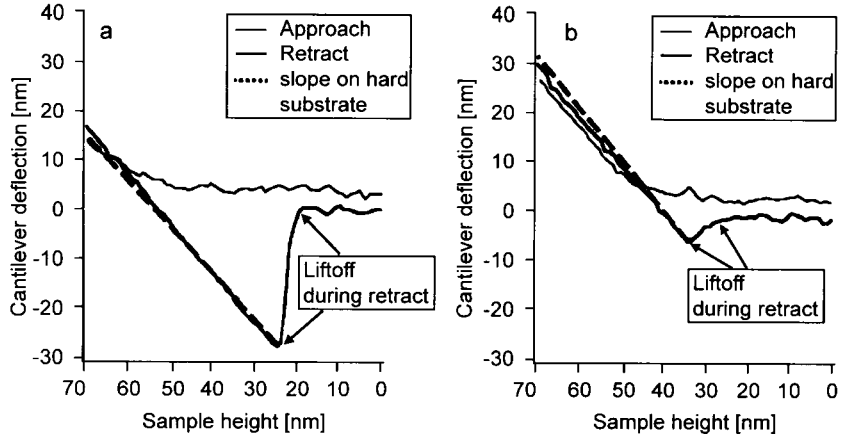


Fig. 19 Force-distance curves on mica and lysozyme. (a) Force curve on the substrate mica. (b) Force curve on an aggregated cluster of lysozyme adsorbed onto mica. Note that the hysteresis during retract is larger on mica, indicating a higher adhesion force. However, liftoff over the lysozyme cluster is gradual in comparison to mica. This effect is probably due to damping of the cantilever by the lysozyme. As a result of elastic indentation, the contact region of the force curve is not just a straight line, but shows some deviation in the case of lysozyme (see fitted dashed contact lines in parts a and b). (Reprinted from Ref. [66] with permission from The American Chemical Society.)

described this initial step as a "fast process". In our work with lysozyme, protein on the surface could be imaged within minutes of exposure to an unstirred protein solution.[45] The second stage is a nonlinear reduction of the adsorption rate until a monolayer plateau is established. During the second stage, the slowing down of adsorption may be explained by lower availability of free adsorption sites on the surface while the protein diffuses, structurally rearranges, and aggregates on the surface. The third stage represents completion of monolayer coverage followed by the fourth stage, where slow multilayer growth begins. The slow growth of a second layer of lysozyme on mica has been reported (Fig. 2).[45] It is thought that the slower kinetics of multilayer adsorption is possibly because the proteins in the bulk solution are largely in their native conformation and do not readily interact with protein that is minimally unfolded at the surface. As the monolayer remains longer on the surface and unfold more, interactions with bulk native protein is likely enhanced.

Currently, the kinetics of protein adsorption at solid surfaces cannot be reliably predicted. This is because of the complicated collection of possible behaviors including irreversibility of protein adsorption, slow conformational changes, and aggregation both within a given layer and between multilayers. The onset of irreversibility can occur in the first and second stages of adsorption. For example, the surface concentration of cytochrome *c* leveled off at a nonzero value when washout began at this second stage.[27] The initial lysozyme clusters observed after 4 min of exposure to mica in Fig. 2 are already irreversibly adsorbed.[45] Subsequent lateral migration and surface aggregation strongly affects the kinetics of adsorption.[27]

Growth Mechanisms: Surface Diffusion/Lateral Migration

Many proteins have been found to diffuse laterally on the surface. Tilton et al.[71,72] have shown using fluorescence recovery after photobleaching (FRAP) that even proteins that are irreversibly adsorbed undergo lateral motion. These workers report surface diffusion coefficient values for bovine serum albumin (BSA) irreversibly adsorbed on poly(methyl methacrylate) (PMMA) and poly(dimethylsiloxane) (PDMS) of $1.2\pm0.3\times10^{-9}$ and $2.6\pm0.1\times10^{-9}$ cm^2/sec, respectively. We explained the clustering of lysozyme molecules on mica as arising from surface diffusion and aggregation of irreversibly adsorbed protein molecules.[45] The irreversible adsorption of lysozyme onto mica probably disturbs the protein conformation sufficiently to expose hydrophobic residues giving rise to surface cluster formation through interprotein aggregation at the surface. Mueller et al.[54] reported a similar mechanism of adsorption, diffusion, and aggregation resulting in clusters of MBP adsorbed onto lipid bilayers. The adsorption patterns shown in Figs. 3 and 4 of IgG and GOx on HOPG also imply surface mobility and aggregation of proteins. Denis et al.[39] attributed the elongated structures of collagen on solids to surface mobility. Collagen formed elongated structures on smooth surfaces exposing $-CH_3$ moieties, but not on rough surfaces or surfaces exposing -OH

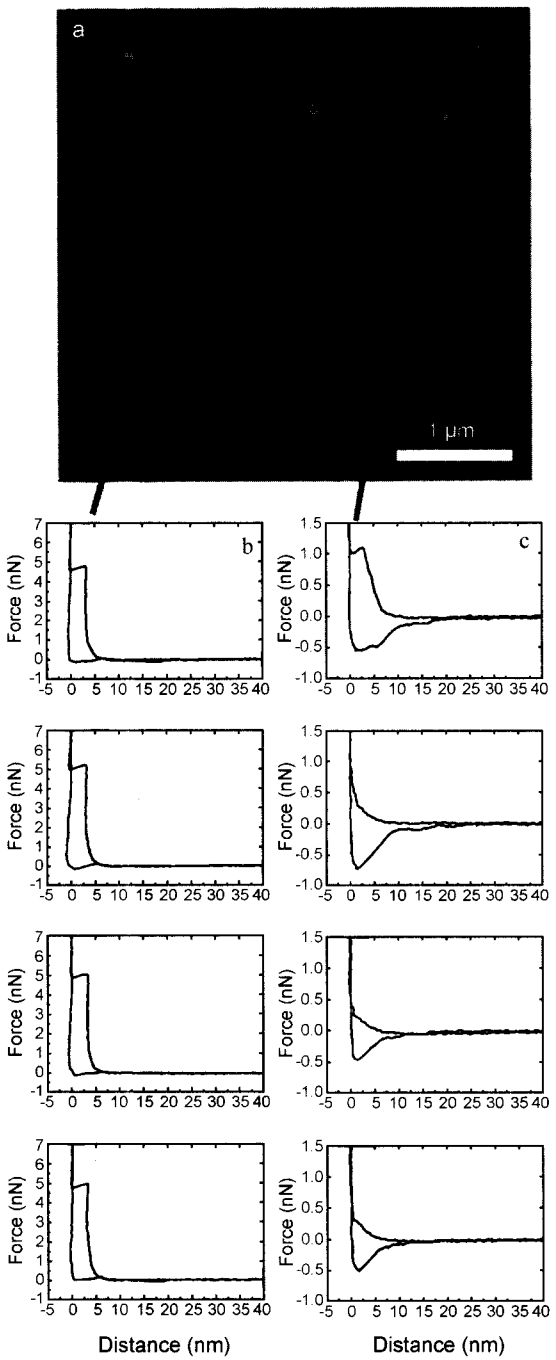


Fig. 20 Sequential, spatially resolved force spectroscopy on an acidic bilayer with adsorbed MBP aggregates. (a) Height image of an acidic lipid bilayer 15 min after injection of 2 μg/mL MBP solution. (b) Column of four consecutive force curves take at the indicated position in the flat area of the image. (c) Column of four consecutive force curves taken on the indicated MBP aggregate. The penetration force was smaller than in b, the jump vanished during the force scans, and an adhesion of approximately 0.5 nN was observed. (Reprinted from Ref. [54] with permission from The American Chemical Society.)

moieties. This was ascribed to the diminished mobility on the rough surfaces. In Hook et al.'s[18] ellipsometry and OWLS studies with serum proteins, it was found that the adsorption saturation mass frequently exceeds the mass predicted by the RSA model. The explanation lies in the ability of protein on the surface to move

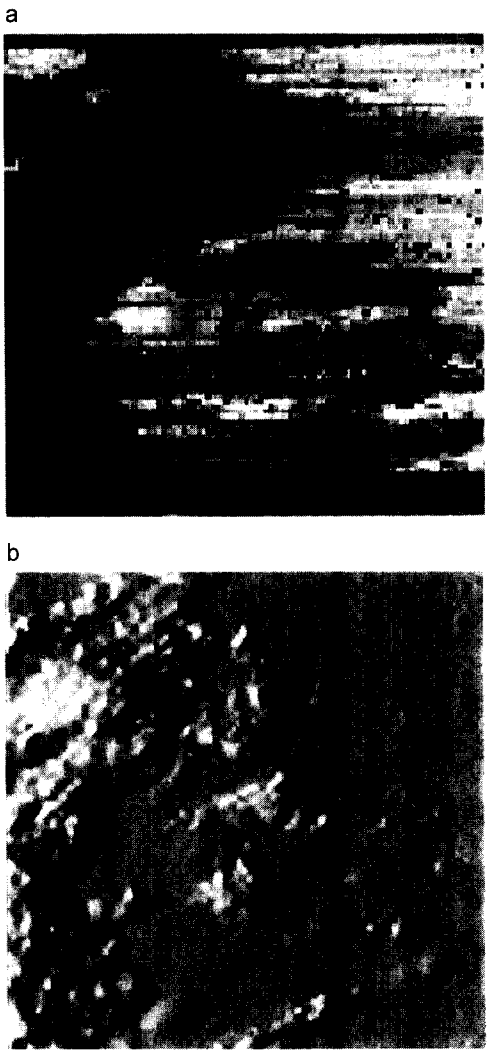


Fig. 21 Force-volume map of BSA adsorbed onto polystyrene. (a) Two-dimensional distribution of the adhesion energy between a BSA-coated probe and a polystyrene surface preadsorbed with BSA, with (b) a simultaneous recorded surface topography (2 μm × 2 μm). Only the left portion of the surface was exposed to the protein solution. The measurements were performed in pH 7 phosphate buffer at the boundary region between the BSA-free and BSA-covered polystyrene surfaces. The adhesion energies are extracted from force–distance curves measured at each sampling point, with brighter shades indicating higher adhesion energy. The topographic image is artificially shadowed for a better presentation. (Reprinted from Ref. [67] with permission from The American Chemical Society.)

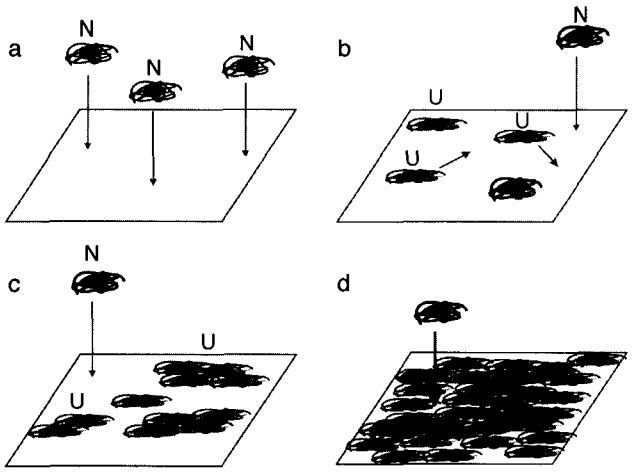


Fig. 22 Generalized illustration of protein adsorption. (a) At early times, protein in its native "N" state rapidly adsorbs onto bare surface. (b) Once adsorbed, protein begins slow unfolding and spreading on the surface (labeled as "U"), rendering them irreversibly adsorbed, but yet mobile on the surface. Additional protein adsorbs onto the surface on available bare surface. (c) Protein aggregates on the surface to form clusters. The rate of protein adsorption onto the surface decreases as less bare surface is available. (d) A complete monolayer is eventually established. Multilayer adsorption and reconfiguration occurs on a slow time scale.

laterally, clearing space for additional protein to adsorb. Calonder et al.[27] reported that the kinetics of additional protein adsorption depends on the amount of time protein has to rearrange and move to form clusters.

Multilayer Formation

Some proteins have the ability to form multilayers. We ascertained that lysozyme forms a second layer, but only after establishing a first monolayer on mica (Fig. 2).[45] The growth of the second layer occurred on a slower time scale than that for the first monolayer, indicating that the adsorption mechanism must be different. This is not surprising considering that once the first monolayer is established, subsequent proteins must adsorb onto a completely different interface. We speculate that the driving force for adsorption on a second layer is exposed, hydrophobic residues resulting from protein conformational changes, occurring in the first layer, that interact with protein in the bulk solution. Hydrophobic interactions may also explain the irreversibility of the second layer of adsorbed lysozyme. Wahlgren et al.[28] and Malmsten[73] likewise reported multilayer adsorption of lysozyme using ellipsometry. Coen et al.[55] also documented multilayer adsorption for protein A onto silicon. As was observed with lysozyme, formation of subsequent layers of protein A did not occur until the previous layer was saturated with protein (Figs. 7 and 8). It was observed that protein in the first monolayer was significantly unfolded (Fig. 7), whereas protein in the second layer was less structurally altered (Fig. 8). Protein in the third layer had no significant structural perturbations.[55] Hook et al.,[18] using OWLS and ellipsometry, found that hemoglobin formed multilayers on titanium oxide surfaces, and Ladam et al.,[30] using SAR, showed that HSA forms multilayers on a positively charged polyallylamine surface.

Nevertheless, not all proteins have the ability to form multilayers. Green et al.[23] found that albumin does not form multilayers on negatively charged polystyrene and neither did HSA.[30] However, multilayer formation can occur over long time scales and most studies do not investigate adsorption beyond a few minutes. Thus it is important to extend the time scale of investigation over several hours or days before eliminating possible irreversible and multilayer adsorption behavior. Also, these phenomena are clearly protein and surface specific.

Protein Structure and Adsorption Behavior

Several studies directly compare different adsorption patterns on different solid surfaces. Immunoglobulin G has been observed to adsorb randomly on HOPG surfaces, forming clusters, whereas GOx prefers to adsorb at HOPG step defects forming "sheets" that eventually grow from these initial nucleation sites (Figs. 3 and 4).[56] Bovine fibrinogen adsorbs onto HOPG in a strand network (Fig. 5), whereas on mica, it adsorbs homogeneously (Fig. 6).[37] This variety of adsorption behavior is most likely related to specific protein structure. The location and the number of hydrophobic and hydrophilic residues can play a large part in the final adsorption pattern. Glucose oxidase is heavily glycosylated and, at pH 7.2, is more negatively charged than is IgG.[74] Thus it is not surprising that GOx nucleates preferentially at higher energy step defects on HOPG rather than on the basal planes. In contrast, IgG is more hydrophobic and can interact more strongly with the hydrophobic surface. Ta et al.[37] explained the discrepancy in the adsorption patterns for BFG on mica and on HOPG based on the native protein structure. The BFG molecule is known to have two distal D domains connected to a central E domain.[75] The D domains are hydrophobic with low structural stability. From each D domain extends an αC domain with both hydrophobic and hydrophilic regions. αC domains are known to have a tendency to interact intra- and intermolecularly with other αC domains. Ta et al.[37] proposed that on HOPG, the D domains hydrophobically bind to the surface while the αC domains interact with other αC domains. In contrast, on negatively charged mica surfaces, the positively charged αC domains likely interact strongly with the surface and are unavailable for intermolecular binding.

The particular protein structure clearly affects the final adsorbed conformation. The von Willebrand factor has a native structure where the termini are extensively disulfide bonded and, therefore, structurally very stable.[76] The termini are also very hydrophilic. The remainder of the molecule is balanced in its hydrophobic and hydrophilic character. On mica, VWF is thought to be more loosely bound; thus, the extended structure of the protein and the high coverages, as shown in Fig. 9, may be a result of a combination of repulsive hydration, electrostatic, and attractive van der Waals interactions.[59] On HOPG, evidence of spreading is shown by its smaller height compared to the native structure (Fig. 10). However, both the amount of lateral spreading and the surface coverage under similar conditions are much smaller than for VWF on mica (compare Figs. 9 and 10). The origin of this difference is not transparent. The authors suggest that on HOPG, "the spreading stresses [of VWF] may be transmitted via the structurally rigid termini causing the VWF to adopt compact structures." Structural stability of the protein is a key factor in determining the extent of conformational change upon adsorption. This was illustrated by Billsten et al.,[77] who measured the extent of change in secondary structure by CD of T4 lysozyme mutants adsorbed onto silica nanoparticles (average diameter ~ 15 nm). Not unexpectedly, it was found that the less structurally sound, native structures lost more secondary conformation upon adsorption.

Protein Structural Changes

Protein molecules adsorbed on a nonspecific surface tend to undergo conformational change. Structural alteration results from the interaction forces between the hydrated surface and the hydrated protein, which disturb the delicate balance of energy and entropy that maintains the protein in a particular conformation in solution. Conformation rearrangement is seen widely in many experiments. Height measurements and antibody-binding experiments by Coen et al.[55] demonstrated that the first layer of protein A adsorbed at the silicon surface is denatured (Fig. 7). Height profiles of both GOx (Fig. 4)[56] and IgG (Fig. 3)[37] on HOPG suggest that the proteins are structurally reconfigured at the surface. Circular dichroism spectroscopy has been used to show that α-chymotrypsin and cutinase change their α-helix content when adsorbed on either Teflon or silica.[14,15] Interleukin-2, upon exposure to silicone rubber tubing for 24 hr, lost 99.5% of its bulk activity. Experimentation with FTIR revealed that the protein adopts a nonnative conformation in which the α-helix, β-sheet, and turn content changes.[7] Sane et al.[8] also found that lysozyme's secondary structure is disturbed after adsorption onto reversed-phase chromatographic supports.

The extent of conformational change is related to many factors including the relative hydrophobicity of the surface, the exposure time of adsorption, and the adsorption density or surface packing of protein. The hydrophobic residues of a protein are usually more concentrated in the core. Upon contact with a hydrophobic surface, the free energy of the system decreases as water molecules that were adjacent to the hydrophobic areas on the protein and on the solid surface are driven out to allow the hydrophobic residues to reside next to the surface. One can imagine that the more hydrophobic a surface is, the greater is the structural rearrangement that will occur. Ta et al.[44] confirmed this assertion in AFM studies of BFG adsorption onto CH_3- and COOH-terminated SAMs. On the hydrophobic -CH_3 surface, the height profile indicates that the protein flattened out, whereas on the more hydrophilic -COOH surface, BFG did not change shape to the same degree. Norde et al.[14,15] established that α-chymotrypsin and cutinase retain some enzymatic activity on hydrophilic silica while losing all activity on hydrophobic Teflon. In a study of BSA adsorption, Norde and Giaconelli[78] showed that irreversible structural changes occurred on polystyrene, but not on silica. Wahlgren and Arnebrant[25] reported that β-lactoglobulin underwent more significant structural changes on methylated silica and polysulfone than on silica. Fibrinogen and albumin were shown to leave larger "footprints" on hydrophobic surfaces as compared to a hydrophilic surface.[21,79] However, proteins do not unfold and denature solely on hydrophobic surfaces. Coen et al.[55] reported a highly denatured monolayer of protein A on hydrophilic silicon.

The structural rearrangement of the protein at a surface is thought to be a "slow" process. Over time, the number of residue contacts with the surface increases, making the surface–protein interaction stronger. The amount of high-density lipoprotein (HDL) that can be desorbed from polyethylene surfaces by surfactant exposure is significantly decreased after 24 hr of storage.[2] In addition to making more and stronger bonds with the surface, the unfolding protein is thought to change its packing structure, possibly aggregating to form intertwining networks. In force–distance curve experiments, Chen et al.[67] discovered that the force required to penetrate through a layer of albumin adsorbed on pyrostyrene increased over time. Friction-force microscopy of BFG on patterned CH_3- and COOH-terminated SAMs revealed that the friction behavior of the adsorbed protein on this surface changes over time, as in Fig. 13.[44] Almeida et al.[70] have shown that the contact angles for monolayers of enolase adsorbed onto both hydrophilic silicon wafers and (aminopropyl)-trimethoxysilane (APS) and hydrophobic PS change with time, indicating structural rearrangements are occurring well after the protein is initially adsorbed to the surface.

The amount of crowding on the surface affects the amount of structural rearrangement that can occur. If the

average nearest neighbor distance is relatively large, then the protein has a greater opportunity to spread. However, if the surface is more crowded, there is only limited space for the protein to spread. Using CD, Norde et al.[14,15] showed that a higher fraction of α-chymotrypsin and cutinase adsorbed on silica remain in their native state if the surface coverage was high. Of course, with time, structural reconfiguration is still likely to occur, but at a slow rate. For BSA on silica and PS, Norde and Giaconnelli[78] reported similar findings. However, Sane et al.[8] showed that the secondary structure of lysozyme is more disturbed on more crowded surfaces. They attributed this finding to lateral interactions between the adsorbed lysozyme.

Irreversible adsorption of protein on a surface is most likely to result from significant structural changes occurring on the surface. As unfolding proceeds, more contacts are created with the surface. In analogy to polymer adsorption, it becomes difficult for the partially denatured protein to be removed. Wahlgren et al.[28] showed that the longer lysozyme was adsorbed on silicon oxide, the larger was the fraction of protein irreversibly adsorbed. This is true for β-lactoglobulin on silicon oxide as well.[26] Given more time to unfold, both the number of contacts with the surface and their strength increases. Foster showed that the irreversibility of albumin adsorption was higher for hydrophobic CH_3-terminated SAMs than for bare silica. Both Wahlgren and Arnebrant[25] and Marsh et al.[26] confirmed this with β-lactoglobulin adsorption onto hydrophilic and hydrophobic surfaces. This was attributed to hydrophobic surfaces being more likely to denature protein to a greater extent than hydrophilic surfaces. However, given the tendency of proteins to reconfigure at long times, the distinction in the amount of protein irreversibly adsorbed on hydrophilic and hydrophobic surfaces is likely to decrease. Wertz and Santore[79] showed that adsorbed protein layers were more susceptible to washout at higher coverage because the increased crowding on the surface limited the amount of solid surface area available per protein molecule.

Fig. 22 shows a general kinetic picture of protein adsorption onto a nonspecific solid surface (i.e., a solid surface without designed specific binding sites for the protein). Particular proteins on particular surfaces may not visit all the listed steps. Initially, in Fig. 22a, protein is in its native conformation (labeled as "N" with a shaded hydrophobic core and a more diffuse hydrophilic corona) in the bulk aqueous environment. Mass transfer processes bring the native protein particle to the surface where it quickly adsorbs onto the mostly empty surface via a combination of van der Waals, electrostatic, and hydrophobic forces. At this stage, interaction forces between the surface and the protein can be weak, and reversible desorption is possible. Once adsorbed at the surface, the protein undergoes structural rearrangement as it is no longer completely surrounded by bulk water. Replacement of surface water by protein segments can be favorable leading to partial unfolding (labeled as "U" in Fig. 22b). The time for and extent of unfolding depends both on surface chemistry (e.g., hydrophobic vs. hydrophilic) and quite importantly on the stability of the native protein. Typically, surface conformation change gives rise to unburying of hydrophobic residues and to lateral "spreading" of the protein adsorbate as more residues can now contact the surface. The amount of spreading depends on the surrounding coverage. Irreversible adsorption and partial denaturation develops at this point. Simple model calculations indicate the conformation change can be extensive creating loops and trains, quite analogous to polymer adsorption from solution.[80,81] Although irreversibly attached to the surface, the individual protein monomers participate in lateral diffusion, as shown in Fig. 22b. Collisions between adsorbate protein molecules permit surface aggregation via interprotein exposed hydrophobic residues. Clusters result, as shown in Fig. 22c, similar to bulk aggregation of partially denatured protein. Upon clustering, the surface aggregates are likely to be immobile because of the increased number of contacts with the surface. As slow reconfiguration and clustering continues, more protein arrives at the surface, but the rate of adsorption slows because of the decreasing availability of bare surface area. It may be possible for protein to adsorb onto already adsorbed protein but this apparently occurs on a longer time scale compared to that onto the unoccupied surface. Typically, the surface fills into a monolayer before multilayers begin forming in Fig. 22d. If the rate of protein arrival to the surface is fast compared to the time scale for lateral diffusion, formation of clusters is bypassed and monolayer coverage is directly reached. Additional protein may adsorb onto the first layer by interaction with hydrophobic residues exposed as a result of the conformation changes induced in the underlying layer. Initially, the multilayer protein might be reversibly adsorbed. However, irreversible incorporation and aggregation is later anticipated. Slow cooperative reconfiguration of the multiplayer protein film may continue for very extended times.

CONCLUSION

Atomic force microscopy is a powerful technique to study protein adsorption. With nanometer lateral resolution of height and force interactions, AFM provides information that optical and spectroscopic techniques cannot. This includes elucidation of the supramolecular assembly of adsorbed protein and of multilayer formation. It can also provide information on of the structure and orientation of the protein as it adsorbs onto the surface. Because of

its excellent lateral resolution, nonuniform surfaces can be studied to compare the adsorption behavior of proteins when presented with various surfaces at the same time. Through its ability to measure forces on a piconewton scale, AFM can measure force interactions between the protein-coated surface and the tip. This gives insight into the forces arising between the protein-coated surface and tip material and the mechanical properties of protein films.

In conjunction with optical and spectroscopic instruments, AFM provides insight into the mechanisms of aqueous protein adsorption onto a solid surface. Electrostatics play a significant role in the amount of protein adsorbed, protein structure, and packing structure. The kinetics of adsorption can be broken down into several stages to reveal kinetic details such as conformation change, surface crowding, monolayer establishment, and multilayer growth. Multilayer growth does not appear to be a universal phenomenon and probably depends on the particular chemical composition and adsorption configuration of a protein. However, we caution that there are only a few long-term (several hours to days) experiments on multilayer formation with a sufficiently wide range of proteins. Once adsorbed on the surface, some proteins exhibit lateral movement, and can form clusters, affecting the total mass adsorbed and the final adsorption patterns. Typically, the particular adsorption patterns observed at the surface can be explained on the basis of the specific protein structure and chemistry. In general, an adsorbed protein undergoes structural perturbation. Structural rearrangement can be minimal or dramatic depending on the system. The extent of unfolding is thought to be related to the hydrophobicity of the surface, although instances of irreversible adsorption on hydrophilic surfaces are legion. Highly unfolded and, therefore, denatured proteins may have many points of contact with the surface and are, thus, more likely to be irreversibly adsorbed. Atomic force microscopy has contributed considerably to the understanding of protein adsorption on nonspecific solid surfaces and will undoubtedly continue to play an important role.

GLOSSARY OF ABBREVIATIONS

AFM	atomic force microscopy
APS	(aminopropyl)-trimethoxysilane
BFG	bovine fibrinogen
BSA	bovine serum albumin
CD	circular dichroism spectroscopy
DETA	trimethoxysilylpropyldiethenetriamine
FOETS	[2-(perfluorooctyl)ethyl]trichlorosilane
FRAP	fluorescence recovery after photobleaching
FTIR	Fourier-transform infrared spectroscopy
GOx	glucose oxidase
HDL	high-density lipoprotein
HEWL	hen egg white lysozyme
HOPG	highly oriented pyrolytic graphite
HSA	human serum albumin
HTS	hexadecyltrichlorosilane
IgG	immunoglobulin G
IgM	immunoglobulin M
MBP	myelin basic protein
MHA	mercaptohexadecanoic acid
ODT	octadecanethiol
OTS	*n*-octadecyltrichlorosilane
OWLS	optical waveguide lightmode spectroscopy
PBS	phosphate-buffered saline
pI	isoelectric point
PDMS	poly(dimethylsiloxane)
PMMA	poly(methyl methacrylate)
PS	polystyrene
QCM	quartz crystal microbalance
RSA	random sequential adsorption
SAM	self-assembled monolayer
SAR	scanning angle reflectometry
SFA	surface force apparatus
SPR	surface plasmon resonance
TIRF	total internal reflection fluorescence
VSFS	vibrational sum frequency spectroscopy
VWF	von Willebrand factor
XPS	X-ray photoelectron spectroscopy

REFERENCES

1. Brange, J.; Langkaer, L. Insulin Structure and Stability. In *Stability and Characterization of Protein and Peptide Drugs*; Wang, Y., Pearlman, R., Eds.; Plenum Press: New York, 1993; 315–348.
2. Feng, M.; Morales, A.B.; Beugeling, T.; Bantjes, A.; Vanderwerf, K.; Gosselink, G.; Degrooth, B.; Greve, J. Adsorption of high density lipoproteins (HDL) on solid surfaces. J. Colloid Interface Sci. **1996**, *177* (2), 364–371.
3. Horbett, T.; Brash, J. Proteins at Interfaces: Current Issues and Future Prospects. In *Proteins at Interfaces—Physiochemical and Biochemical Studies*; Brash, J., Horbett, T., Eds.; American Chemical Society: Washington, DC, 1987; 1–33.
4. Andrade, J.D.; Hlady, V. Protein adsorption and materials biocompatibility—A tutorial review and suggested hypotheses. Adv. Polym. Sci. **1986**, *79*, 1–63.
5. Ivarsson, B.; Lundstrom, I. Physical characterization of protein adsorption on metal and metaloxide surfaces. Crit. Rev. Biocompat. **1986**, *2* (1), 1–96.
6. Norde, W. Adsorption of proteins from solution at

the solid–liquid interface. Adv. Colloid Interface Sci. **1986**, *25* (4), 267–340.
7. Tzannis, S.T.; Hrushesky, W.J.M.; Wood, P.A.; Przybycien, T.M. Adsorption of a formulated protein on a drug delivery device surface. J. Colloid Interface Sci. **1997**, *189* (2), 216–228.
8. Sane, S.U.; Cramer, S.M.; Przybycien, T.M. Protein structure perturbations on chromatographic surfaces. J. Chromatogr., A **1999**, *849* (1), 149–159.
9. Benmakroha, Y.; Zhang, S.; Rolfe, P. Haemocompatibility of invasive sensors. Med. Biol. Eng. Comput. **1995**, *33* (6), 811–821.
10. Tighe, B.; Franklin, V. Lens Deposition and Spoliation. In *The Eye in Contact Lens Wear*; Larke, J.R., Ed.; Butterworth-Heinemann: Oxford, England, 1997; 49–100.
11. Kim, G.; Gurau, M.; Kim, J.; Cremer, P.S. Investigations of lysozyme adsorption at the air/water and quartz/water interfaces by vibrational sum frequency spectroscopy. Langmuir **2002**, *18* (7), 2807–2811.
12. Claesson, P.M.; Blomberg, E.; Froberg, J.C.; Nylander, T.; Arnebrant, T. Protein interactions at solid surfaces. Adv. Colloid Interface Sci. **1995**, *57*, 161–227.
13. Blomberg, E.; Claesson, P.M.; Froberg, J.C.; Tilton, R.D. Interaction between adsorbed layers of lysozyme studied with the surface force technique. Langmuir **1994**, *10* (7), 2325–2334.
14. Zoungrana, T.; Findenegg, G.H.; Norde, W. Structure, stability, and activity of adsorbed enzymes. J. Colloid Interface Sci. **1997**, *190* (2), 437–448.
15. Norde, W.; Zoungrana, T. Surface-induced changes in the structure and activity of enzymes physically immobilized at solid/liquid interfaces. Biotechnol. Appl. Biochem. **1998**, *28*, 133–143.
16. Sane, S.U.; Cramer, S.M.; Przybycien, T.M. A holistic approach to protein secondary structure characterization using amide I band Raman spectroscopy. Anal. Biochem. **1999**, *269* (2) 255–272.
17. Otzen, D.E.; Oliveberg, M.; Hook, F. Adsorption of a small protein to a methyl-terminated hydrophobic surface: Effect of protein-folding thermodynamics and kinetics. Colloids Surf., B Biointerfaces **2003**, *29* (1), 67–73.
18. Hook, F.; Voros, J.; Rodahl, M.; Kurrat, R.; Boni, P.; Ramsden, J.J.; Textor, M.; Spencer, N.D.; Tengvall, P.; Gold, J.; Kasemo, B. A comparative study of protein adsorption on titanium oxide surfaces using in situ ellipsometry, optical waveguide lightmode spectroscopy, and quartz crystal microbalance/dissipation. Colloids Surf., B Biointerfaces **2002**, *24* (2), 155–170.
19. Wertz, C.F.; Santore, M.M. Adsorption and relaxation kinetics of albumin and fibrinogen on hydrophobic surfaces: Single-species and competitive behavior. Langmuir **1999**, *15* (26), 8884–8894.
20. Buijs, J.; Britt, D.W.; Hlady, V. Human growth hormone adsorption kinetics and conformation on self-assembled monolayers. Langmuir **1998**, *14* (2), 335–341.
21. Wertz, C.F.; Santore, M.M. Effect of surface hydrophobicity on adsorption and relaxation kinetics of albumin and fibrinogen: Single-species and competitive behavior. Langmuir **2001**, *17* (10), 3006–3016.
22. Robeson, J.L.; Tilton, R.D. Spontaneous reconfiguration of adsorbed lysozyme layers observed by total internal reflection fluorescence with a pH-sensitive fluorophore. Langmuir **1996**, *12* (25), 6104–6113.
23. Green, R.J.; Davies, J.; Davies, M.C.; Roberts, C.J.; Tendler, S.J.B. Surface plasmon resonance for real time in situ analysis of protein adsorption to polymer surfaces. Biomaterials **1997**, *18* (5), 405–413.
24. Malmsten, M.; Lassen, B.; Westin, J.; Golander, C.G.; Larsson, R.; Nilsson, U.R. Adsorption of complement protein c3 at polymer surfaces. J. Colloid Interface Sci. **1996**, *179* (1), 163–172.
25. Wahlgren, M.; Arnebrant, T. Adsorption of β-lactoglobulin onto silica, methylated silica, and polysulfone. J. Colloid Interface Sci. **1990**, *136* (1), 259–265.
26. Marsh, R.J.; Jones, R.A.L.; Sferrazza, M. Adsorption and displacement of globular protein on hydrophilic and hydrophobic surfaces. Colloids Surf., B Biointerfaces **2002**, *23* (1), 31–42.
27. Calonder, C.; Tie, Y.; van Tassel, P.R. History dependence of protein adsorption kinetics. Proc. Natl. Acad. Sci. U. S. A. **2001**, *98* (19), 10664–10669.
28. Wahlgren, M.; Arnebrant, T.; Lundstrom, I. The adsorption of lysozyme to hydrophilic silicon oxide surfaces—Comparison between experimental data and models for adsorption kinetics. J. Colloid Interface Sci. **1995**, *175* (2), 506–514.
29. Ramsden, J.J.; Prenosil, J.E. Effect of ionic strength on protein adsorption kinetics. J. Phys. Chem. **1994**, *98* (20), 5376–5381.
30. Ladam, G.; Gergely, C.; Senger, B.; Decher, G.; Voegel, J.-C.; Schaaf, P.; Cuisinier, F.J.G. Protein interactions with polyelectrolyte multilayers: Interactions between human serum albumin and polystyrene sulfonate/polyallylamine multilayers. Biomacromolecules **2000**, *1* (4), 674–687.
31. Fritz, M.; Radmacher, M.; Cleveland, J.P.; Allersma, M.W.; Stewart, R.J.; Gieselmann, R.; Janmey, P.; Schmidt, C.F.; Hansma, P.K. Imaging

obular and filamentous proteins in physiological buffer solutions with tapping mode atomic force microscopy. Langmuir **1995**, *11* (9), 3529–3535.
32. Bustamante, C.; Rivetti, C. Visualizing protein–nucleic acid interactions on a large scale with the scanning force microscope. Annu. Rev. Biophys. Biomol. Struct. **1996**, *25*, 395–429.
33. Bergkvist, M.; Carlsson, J.; Karlsson, T.; Oscarsson, S. TM–AFM threshold analysis of macromolecular orientation: A study of the orientation of IgG and IgE on mica surfaces. J. Colloid Interface Sci. **1998**, *206* (2), 475–481.
34. Muller, D.J.; Sass, H.-J.; Muller, S.A.; Buldt, G.; Engel, A. Surface structures of native bacteriorhodopsin depend on the molecular packing arrangement in the membrane. J. Mol. Biol. **1999**, *285* (5), 1903–1909.
35. Muller, D.J.; Engel, A. pH and voltage induced structural changes of porin ompf explain channel closure. J. Mol. Biol. **1999**, *285* (4), 1347–1351.
36. Smith, D.A.; Connell, S.D.; Robinson, C.; Kirkham, J. Chemical force microscopy: Applications in surface characterization of natural hydroxyapatite. Anal. Chim. Acta **2003**, *479* (1), 39–57.
37. Ta, T.C.; Sykes, M.T.; McDermott, M.T. Real-time observation of plasma protein film formation on well-defined surfaces with scanning force microscopy. Langmuir **1998**, *14* (9), 2435–2443.
38. McMaster, T.J.; Miles, M.J.; Shewry, P.R.; Tatham, A.S. In situ surface adsorption of the protein c hordein using atomic force microscopy. Langmuir **2000**, *16* (4), 1463–1468.
39. Denis, F.A.; Hanarp, P.; Sutherland, D.S.; Gold, J.; Mustin, C.; Rouxhet, P.G.; Dufrene, Y.F. Protein adsorption on model surfaces with controlled nanotopography and chemistry. Langmuir **2002**, *18* (3), 819–828.
40. Johnson, C.A.; Yuan, Y.; Lenhoff, A.M. Adsorbed layers of ferritin at solid and fluid interfaces studied by atomic force microscopy. J. Colloid Interface Sci. **2000**, *223* (2), 261–272.
41. Lin, J.N.; Drake, B.; Lea, A.S.; Hansma, P.K.; Andrade, J.D. Direct observation of immunoglobulin adsorption dynamics using the atomic force microscope. Langmuir **1990**, *6* (2), 509–511.
42. Lea, A.S.; Pungor, A.; Hlady, V.; Andrade, J.D.; Herron, J.N.; Voss, E.W. Manipulation of proteins on mica by atomic force microscopy. Langmuir **1992**, *8* (1), 68–73.
43. Gunning, A.P.; Mackie, A.R.; Wilde, P.J.; Morris, V.J. In situ observation of the surfactant-induced displacement of protein from a graphite surface by atomic force microscopy. Langmuir **1999**, *15* (13), 4636–4640.
44. Ta, T.C.; McDermott, M.T. Mapping interfacial chemistry induced variations in protein adsorption with scanning force microscopy. Anal. Chem. **2000**, *72* (11), 2627–2634.
45. Anderson, R.E.; Pande, V.S.; Radke, C.J. Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface. J. Chem. Phys. **2000**, *112* (20), 9167–9185.
46. Howald, L.; Haefke, H.; Luthi, R.; Meyer, E.; Gerth, G.; Rudin, H.; Guntherodt, H.-J. Ultrahigh-vacuum scanning force microscopy—Atomic scale resolution at monatomic cleavage steps. Phys. Rev., B **1994**, *49* (8), 5651–5656.
47. Grutter, P.; Zimmermann-Edling, W.; Brodbeck, D. Tip artifacts of microfabricated force sensors for atomic force microscopy. Appl. Phys. Lett. **1992**, *60* (22), 2741–2743.
48. Sheller, N.B.; Petrash, S.; Foster, M.D. Atomic force microscopy and x-ray reflectivity studies of albumin adsorbed onto self-assembled monolayers of hexadecyltrichlorosilane. Langmuir **1998**, *14* (16), 4535–4544.
49. Villarrubia, J.S. Morphological estimation of tip geometry for scanned probe microscopy. Surf. Sci. **1994**, *321* (3), 287–300.
50. Villarrubia, J.S. Algorithms for scanned probe microscope image simulation, surface reconstruction, and tip estimation. J. Res. Natl. Inst. Stand. **1997**, *102* (4), 425–454.
51. Takahara, A.; Ge, S.R.; Kojio, K.; Kajiyama, T. In situ atomic force microscopic observation of albumin adsorption onto phase-separated organosilane monolayer surface. J. Biomater. Sci., Polym. Ed. **2000**, *11* (1), 111–120.
52. Takahara, A.; Hara, Y.; Kojio, K.; Kajiyama, T. Scanning force microscopic study of protein adsorption on the surface of organosilane monolayers prepared by the Langmuir–Blodgett method. Macromol. Symp. **2001**, *167*, 271–284.
53. Takahara, A.; Hara, Y.; Kojio, K.; Kajiyama, T. Plasma protein adsorption behavior onto the surface of phase-separated organosilane monolayers on the basis of scanning force microscopy. Colloids Surf., B Biointerfaces **2002**, *23* (2–3), 141–152.
54. Mueller, H.; Butt, H.J.; Bamberg, E. Adsorption of membrane-associated proteins to lipid bilayers studied with an atomic force microscope: Myelin basic protein and cytochrome *c*. J. Phys. Chem., B **2000**, *104* (18), 4552–4559.
55. Coen, M.C.; Lehmann, R.; Groning, P.; Bielmann, M.; Galli, C.; Schlapbach, L. Adsorption and bioactivity of protein A on silicon surfaces studied by AFM and XPS. J. Colloid Interface Sci. **2001**, *233* (2), 180–189.

56. Cullen, D.C.; Lowe, C.R. AFM studies of protein adsorption: 1. Time-resolved protein adsorption to highly oriented pyrolytic graphite. J. Colloid Interface Sci. **1994**, *166* (1), 102–108.
57. Hecht, H.J.; Kalisz, H.M.; Hendle, J.; Schmid, R.D.; Schomburg, D. Crystal structure of glucose-oxidase from *Aspergillus niger* refined at 2.3 angstrom resolution. J. Mol. Biol. **1993**, *229* (1), 153–172.
58. Silverton, E.W.; Navia, M.A.; Davies, D.R. 3-Dimensional structure of an intact human immunoglobulin. Proc. Natl. Acad. Sci. U. S. A. **1977**, *74* (11), 5140–5144.
59. Raghavachari, M.; Tsai, H.M.; Kottke-Marchant, K.; Marchant, R.E. Surface dependent structures of von Willebrand factor observed by AFM under aqueous conditions. Colloids Surf., B Biointerfaces **2000**, *19* (4), 315–324.
60. Blake, C.C.F.; Koenig, D.F.; Mair, G.A.; North, A.C.T.; Phillips, D.C.; Sarma, V.R. Structure of hen egg-white lysozyme—A 3-dimensional Fourier synthesis at 2 angstrom resolution. Nature **1965**, *206* (4986), 757.
61. Field, J.M.; Tatham, A.S.; Baker, A.M.; Shewry, P.R. The structure of c hordein. FEBS Lett. **1986**, *200* (1), 76–80.
62. Galli, C.; Coen, M.C.; Hauert, R.; Katanaev, V.L.; Groning, P.; Schlapbach, L. Creation of nanostructures to study the topographical dependency of protein adsorption. Colloids Surf., B Biointerfaces **2002**, *26* (3), 255–267.
63. Clark, P.; Connolly, P.; Curtis, A.S.G.; Dow, J.A.T.; Wilkinson, C.D.W. Topographical control of cell behavior: 2. Multiple grooved substrata. Development **1990**, *108* (4), 635–644.
64. Clark, P.; Connolly, P.; Curtis, A.S.G.; Dow, J.A.T.; Wilkinson, C.D.W. Cell guidance by ultrafine topography in vitro. J. Cell Sci. **1991**, *99*, 73–77.
65. Curtis, A.; Wilkinson, C. Topographical control of cells. Biomaterials **1997**, *18* (24), 1573–1583.
66. Radmacher, M.; Fritz, M.; Cleveland, J.P.; Walters, D.A.; Hansma, P.K. Imaging adhesion forces and elasticity of lysozyme adsorbed on mica with the atomic force microscope. Langmuir **1994**, *10* (10), 3809–3814.
67. Chen, X.; Davies, M.C.; Roberts, C.J.; Tendler, C.J.B.; Williams, P.M. Recognition of protein adsorption onto polymer surfaces by scanning force microscopy and probe surface adhesion measurements with protein-coated probes. Langmuir **1997**, *13* (15), 4106–4111.
68. Patel, N.; Davies, M.C.; Heaton, R.J.; Roberts, C.J.; Tendler, S.J.B.; Williams, P.M. A scanning probe microscopy study of the physisorption and chemisorption of protein molecules onto carboxylate terminated self-assembled monolayers. Appl. Phys., A **1998**, *66* (PT1 SUPPS), S569–S574.
69. Lee, S.H.; Ruckenstein, E. Adsorption of proteins onto polymeric surfaces of different hydrophilicities—A case-study with bovine serum-albumin. J. Colloid Interface Sci. **1988**, *125* (2), 365–379.
70. Almeida, A.T.; Salvadori, M.C.; Petri, D.F.S. Enolase adsorption onto hydrophobic and hydrophilic solid substrates. Langmuir **2002**, *18* (18), 6914–6920.
71. Tilton, R.D.; Gast, A.P.; Robertson, C.R. Surface diffusion of interacting proteins—Effect of concentration on the lateral mobility of adsorbed bovine serum albumin. Biophys. J. **1990**, *58* (5), 1321–1326.
72. Tilton, R.D.; Robertson, C.R.; Gast, A.P. Lateral diffusion of bovine serum albumin adsorbed at the solid liquid interface. J. Colloid Interface Sci. **1990**, *137* (1), 192–203.
73. Malmsten, M. Ellipsometry studies of protein layers adsorbed at hydrophobic surfaces. J. Colloid Interface Sci. **1994**, *166* (2), 333–342.
74. Wilson, R.; Turner, A.P.F. Glucose-oxidase—An ideal enzyme. Biosens. Bioelectron. **1992**, *7* (3), 165–185.
75. Feng, L.; Andrade, J.D. *Proteins at Interfaces II: Fundamentals and Applications*; Brash, J.L., Horbett, T.A., Eds.; American Chemical Society: Washington, DC, 1995; Chapter 5.
76. Titani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.A.; Chopek, M.W.; Sadler, J.E.; Fujikawa, K. Amino-acid-sequences of human von Willebrand-factor. Biochemistry **1986**, *25* (11), 3171–3184.
77. Billsten, P.; Wahlgren, M.; Arnebrant, T.; McGuire, J.; Elwing, H. Structural changes of t4 lysozyme upon adsorption to silica nanoparticles measured by circular dichroism. J. Colloid Interface Sci. **1995**, *175* (1), 77–82.
78. Norde, W.; Giacomelli, C.E. BSA structural changes during homomolecular exchange between the adsorbed and the dissolved states. J. Biotechnol. **2000**, *79* (3), 259–268.
79. Wertz, C.F.; Santore, M.M. Fibrinogen adsorption on hydrophilic and hydrophobic surfaces: Geometrical and energetic aspects of interfacial relaxations. Langmuir **2002**, *18* (3), 706–715.
80. Anderson, R.E. *Dynamic Monte Carlo Simulation of Protein Interaction with an Oil/Water Interface*; University of California: Berkeley, CA, 2000; Masters thesis.
81. Castells, V.; Yang, S.X.; Van Tassel, P.R. Surface-induced conformational changes in lattice model proteins by Monte Carlo simulation. Phys. Rev., E **2002**, *65* (3). Art. No. 031912.

Protein Nanotubes as Building Blocks

P

Hiroshi Matsui
The City University of New York, Hunter College, New York, New York, U.S.A.

INTRODUCTION

Nanometer scale self-assemblies of biological molecules, such as protein nanotubes, play important roles in nature to regulate biological activities. It is a natural progress to synthesize protein nanotubes in order to mimic those biological functions and apply for medical applications such as drug delivery and medical imaging. Although it appears to be unlikely, peptide nanotubes are also applied to non-biological fields such as electronics, sensors, catalysis, and optics because those smart, selective peptide functions can be applied in various device fabrication processes. Those new types of building blocks with smart biological functions may offer more economical, reproducible, accurate, and simpler methods in device fabrications. In this article, various examples of natural and synthetic protein nanotubes and their functionalizations for particular applications are discussed.[1–32]

OVERVIEW

Protein nanotubes are tubular assemblies of proteins. The production of protein nanotubes can be synthetic or natural. They are generally produced via their direct self-assemblies into tubular structures[33] or their self-assemblies on template nanotubes.[34–37] While the direct nanotube assemblies from protein monomers without templates are normally achieved in vitro,[38,39] there are some examples that protein nanotubes are self-assembled in solutions in the presence of divalent metal ions.[40] With the aid of surfactants, some proteins were also demonstrated to be assembled as a tubular form.[41] Molecular self-assembly is robust and reproducible under the controlled experimental conditions.[42,43] For example, protein nanotubes are grown in solution via self-assembly with high reproducibility and monodispersity.[40] Thermodynamics controls chemical structures and dimensions of self-assembled nanotubes[44] and it can be determined by external experimental conditions such as pH, temperature, solvents, and pressure.[45] Protein nanotubes such as virus, phages, and microtubules are also produced via natural self-assembly with high reproducibility and monodispersity inside cells. The chemical structures of natural protein nanotubes can be modified via mutations, which gives flexibility to design protein nanotubes.[46] Therefore the use of self-assemblies either synthetically or naturally has an advantage to mass-produce monodisperse and chemically identical nanotubes. Because protein nanotubes are water-soluble, it makes their size separation easier if necessary.

While the self-assembly nature offers some interesting features in protein nanotubes, they may not display useful physical properties as shown in metal and semiconductor nanowires for particular devices such as electronics and sensors. Then what makes protein nanotubes special compared with other inorganic nanotubes/nanowires? I believe that features of protein nanotubes distinguished from other nanotubes are their flexibility in the functionality and their molecular recognition. All proteins have some functions and those functions can be used to introduce new properties into protein nanotubes, desired for particular building blocks in specific nanodevices. For example, proteins can mineralize various metals and semiconductors selectively via their molecular recognition and biomineralization, which are routinely operated in living systems.[47,48] Positioning of protein nanotubes can be achieved through biological recognitions of proteins toward their complementary ligands.

In this article, I would like to categorize the features of protein nanotubes into four characteristic functions: 1) coating, 2) arrangement, 3) sensing, and 4) mechanical motion. In the coating section, biomineralization of protein nanotubes is discussed. In the placement section, organization and location-specific immobilization of protein nanotubes are explained via biological recognitions such as inorganic-protein and protein–protein interactions. Physical property changes of protein nanotubes via biological recognition-induced adsorption of target molecules can be used for their detections and quantitative analysis, as explained in the "Sensing" section. In the "Mechanical motion" section, motor functions of protein nanotubes by converting chemical energy into mechanical force are summarized.

COATING

To apply protein nanotubes as building blocks for various nanodevices, functionalizations of protein nanotubes via

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013709

coatings are very effective. Choices of proteins are dependent upon what types of materials to be coated. Tobacco mosaic virus (TMV) was used to mineralize CdS, PbS, silica, and iron oxides on the outer surfaces of TMV.[49,50] Tobacco mosaic virus, a hollow protein nanotube, is a stable viron with 18 nm in diameter and 300 nm in length, self-assembled from the protein monomers. The repeated protein subunits on the nanotubes offer the nucleation sites to produce highly crystalline semiconductor coatings. Fig. 1a shows CdS nanotubes fabricated by exposing a buffered suspension of TMV in 10 mM $CdCl_2$ to H_2S gas for up to 6 hr. High-resolution transmission electron micrograph (HRTEM) of the CdS nanotubes in Fig. 1b indicates that the CdS coating consists of aggregates of crystalline CdS nanoparticles with 5 nm in diameter, whose fringe space matches a (111) lattice plane of the zinc-blende structure. Modified TMV nanotubes at the C-terminus via a single site mutation were also synthesized as template nanotubes to improve their selectable chemospecificity.[46] The modified TMV nanotubes have lysine units on the outer surfaces of the nanotubes and these sites can be used to immobilize functional molecules via *N*-hydroxy succinamide chemistry (Fig. 2). This mutant TMV nanotube functionalized with flurophor selectively mineralized Cu via photoreduction of Cu ions. Another example of biomineralization of protein nanotubes is to produce magnetic nanofibers by using rod-shaped bacteriums as templates.[51] A single *bacillus subtilis* bacterium is an organism with cells of approximately 0.7 μm in diameter and up to 4 μm in length. Two types of iron oxide nanoparticles, magnetite (Fe_3O_4) and maghaemite (γ-Fe_2O_3), were mineralized on the bacterial threads. The magnetization data showed that the magnetic field-dependent switching behavior was markedly sharp when single domains of iron oxide nanoparticles were aligned to the easy axes of the nanotubes. This alignment was achieved when the magnetic nanotube synthesis was conducted within an external magnetic field.

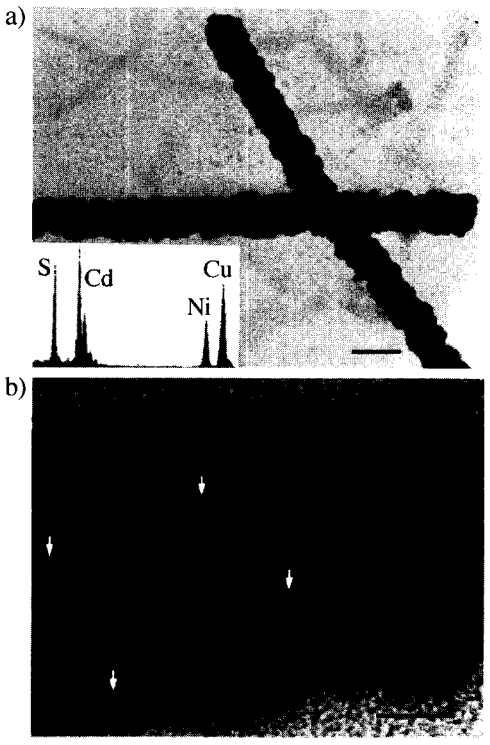


Fig. 1 TEM images of (a) CdS-coated TMV nanotubes, scale bar=50 nm, inset: corresponding EDX spectrum (Ni and Cu peaks arise from TEM grid and sample holder). (b) CdS nanocrystals (white arrows) on the TMV nanotube in high magnification, scale bar=10 nm. (From Ref. [49]. Copyright (1999) Wiley-VCH Verlag GmbH.)

Microtubules, self-assembled from cytoskeletal proteins in vitro, were used as template nanotubes to coat Pd.[38] The subunits of purified porcine brain tubulin were cross-linked by glutaric dialdehyde and this process yielded stabilized microtubules, whose diameter is 4–5 nm and length is about 8 nm. This resulting microtubule remained intact up to 90°C for approximately 1 hr in the pH range from 4 to 9. The improved microtubule mineralized Pd with trisodium citrate or dimethylamine borane at 90°C. The trisodium citrate produced smaller Pd nanocrystals with a mean diameter of 1.9 nm and the dimethylamine borane produced larger Pd nanocrystals with a mean diameter of 3.1 nm (Fig. 3a). HRTEM micrograph of the highly monodisperse Pd nanocrystals on the nanotubes imaged the (111) plane of a pure metallic Pd phase (Fig. 3b). Pd nanocrystals were found to grow on both α- and β-tubulin subunits of the protein nanotubes and the histidine residues likely serve as nucleation sites for the Pd biomineralization on the nanotubes.

Improved metal nanocrystal coatings were demonstrated on the biological nanotubes incorporating sequenced peptides that can mineralize specific metals via biological recognitions.[52] This sequence consisting of histidine and alanine residues not only produced monodisperse Au nanocrystals on the nanotube surfaces but also controlled the density and the size of nanocrystals by tuning the sequenced peptide conformations (Fig. 4).[53] The histidine residue captured Au ions ($ClAuPMe_3$) and the reduction of those trapped Au ions led to monodisperse nanocrystal formation with an average diameter of 5 nm on the nanotubes. The change in experimental conditions such as pHs and ion concentrations in the nanocrystal growth solution induced the ion-peptide conformation change and this structural transformation controlled the size and the distribution of Au nanocrystals on the nanotubes. It should be noted that metallic nanocrystals with a diameter below 6 nm are in the size domain to

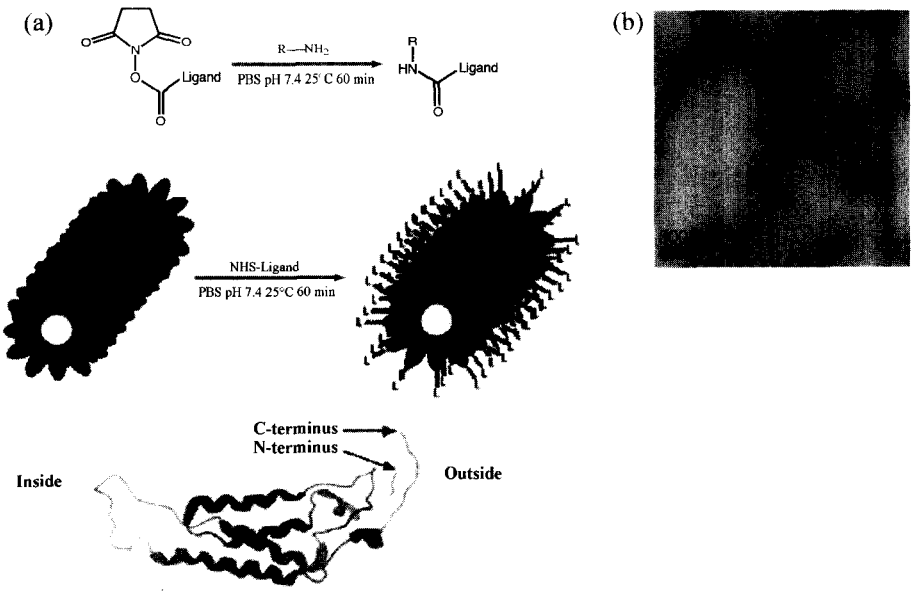


Fig. 2 (a) The upper panel: the facile NHS chemical reaction; the middle panel: the structure of TMV nanotube and a schematic to introduce ligands on the TMV nanotube surface via the NHS chemistry; the lower panel: the structure of the protein monomer for the TMV nanotube and the relative positions of the monomer in the TMV nanotube. (b) Electron micrograph of m-TMV nanotubes. (From Ref. [46]. Copyright (2002) Institute of Physics Publishing.) (*View this art in color at www.dekker.com.*)

observe significant conductivity change by changing the packing density.[54] Therefore this system may be useful as a conductivity-tunable building block by controlling the nanocrystal size and the distribution.

ARRANGEMENT

There are several techniques to align nanotubes by using electronic fields, magnetic fields, microfluidics, and hydrophobic interactions.[26–30] But when multiple types of nanotubes are necessary to be aligned respectively in more complex device configurations, the above methods may not be sufficient. For example, consider fabricating a nanodevice whose configuration has two nanotubes crossing perpendicular to each other (Fig. 5). In this configuration, when one nanotube is coated by a metal and the other nanotube is coated by a semiconductor, this nanodevice will function as a switch in memory circuits for reading and writing information.[6,55] Whereas these

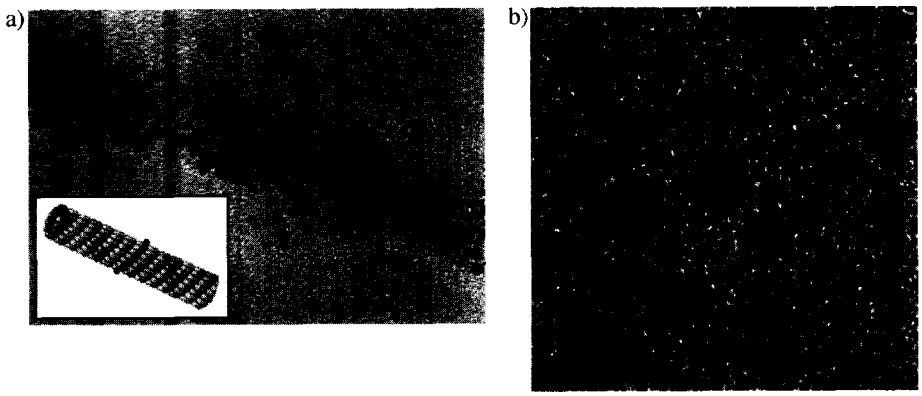


Fig. 3 (a) TEM image of palladium nanoparticles immobilized on a microtubule, scale bar=20 nm. (b) HRTEM image of palladium nanoparticles on the microtubule, scale bar=5 nm. (From Ref. [38]. Copyright (2002) Wiley-VCH Verlag GmbH.)

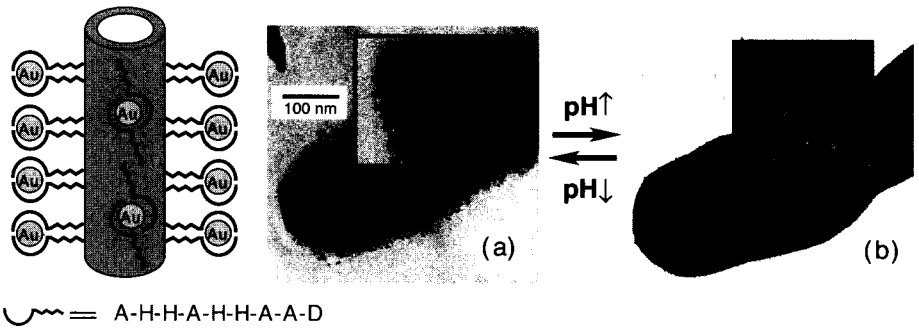


Fig. 4 TEM images of Au nanocrystals regulated by the sequenced histidine-rich peptides on the nanotubes. Au nanocrystals were grown by reducing the Au ion-nanotube solution after incubating Au ions for 10 days at (a) pH=8.5, and (b) pH=11.5. Insets show the TEM images in higher magnification. (From Refs. [52] and [53].) (*View this art in color at www.dekker.com.*)

device elements can be switched between well-defined on-and-off states by transiently charging the coated nanotubes to produce attractive and repulsive electrostatic forces,[55] aligning two or more types of nanotubes in desired directions precisely is not an easy task by the techniques described above.

By functionalizing circuit elements and connecting wires with biomolecular recognition of complementary connecting units while in solution, we can mimic biological systems, in which organic/inorganic nanoscale building blocks are routinely and precisely turned into complex structures for biological function with almost perfect reproducibility. In theory, protein nanotubes, which can recognize and selectively bind a well-defined region on complimentary protein-patterned substrates, can be used as building blocks to assemble three-dimensional nanoscale architectures by placing them at the uniquely defined positions. As shown in Fig. 5, an antigen 1 nanotube can be aligned on antibody 1 areas while an antigen 2 nanotube recognizes and immobilizes on antibody 2 surfaces. Therefore the immobilization of protein nanotubes onto the antibody self-assembled monolayers (SAMs) will be "programmed" (i.e., protein nanotubes recognize binding sites and attach automatically) in aqueous solution without any complicated multistep fabrication procedures. To demonstrate the feasibility of this scheme, the streptavidin-coated nanotubes were dispersed on biotin-incorporated SAMs in solution, and their selective immobilization onto the biotin-incorporated SAMs was observed after 1 day (Fig. 6).[56] Protein-functionalized nanotubes were positioned to interconnect patterned complimentary protein-SAM/Au surfaces via protein–protein interactions; however, uniform coverage of proteins on template nanotubes also caused the nanotube aggregations on the complimentary protein-SAM surfaces.[56] One solution to produce the nanotube-bridge configurations with minimum nanotube aggregations could be selective protein immobilization at the ends of nanotubes. We recently demonstrate that this location-specific protein immobilization on nanotubes was established by using Au nanocrystal masks.[57] As a proof-of-principle, avidin

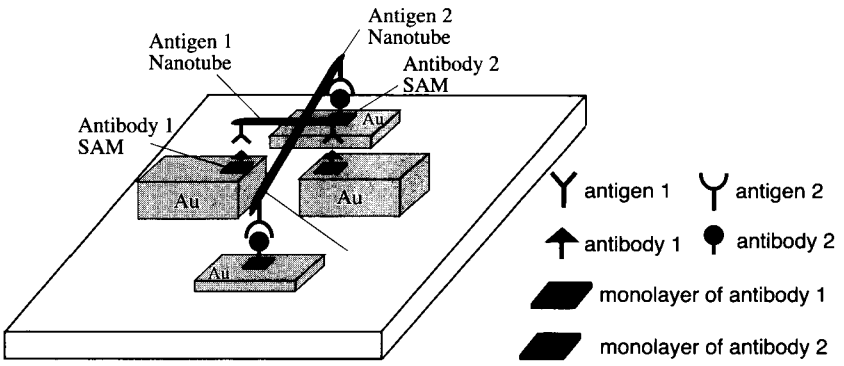


Fig. 5 Illustration of proposed scheme to assemble multiple protein nanotubes into device configurations via biological recognitions between the protein nanotubes and the complimentary protein-patterned surfaces. (*View this art in color at www.dekker.com.*)

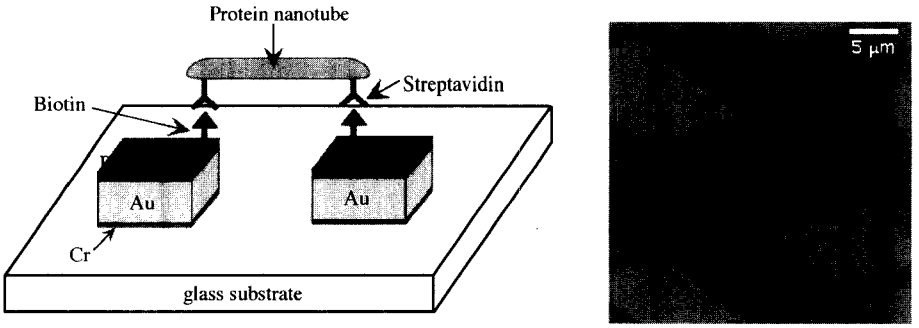


Fig. 6 Illustration and SEM image of a streptavidin-functionalized nanotube immobilized onto the complimentary biotin SAMs/Au substrates. (From Ref. [56].) (*View this art in color at www.dekker.com.*)

was used to be immobilized only at the ends of peptide nanotubes with the nanotube end-fabrication method, as outlined in Fig. 7. The approach involves thiolation of the nanotube sidewalls [step (i)] and Au nanocrystal coating to mask the sidewalls of nanotubes [step (ii)], followed by incubating nanotubes with sulforhodamine-labeled avidin [step (iii)]. After avidin molecules were immobilized at the ends (with no Au nanocrystals) and the sidewalls (with Au nanocrystals) of the nanotubes, the Au nanocrystals on the sidewalls were chemically etched and the proteins at the ends remained attached without denaturing [step (iv)]. The chemical etching of the Au nanocrystal masks on the sidewalls of nanotubes removed avidin molecules from the sidewalls; however, avidin molecules at the nanotube

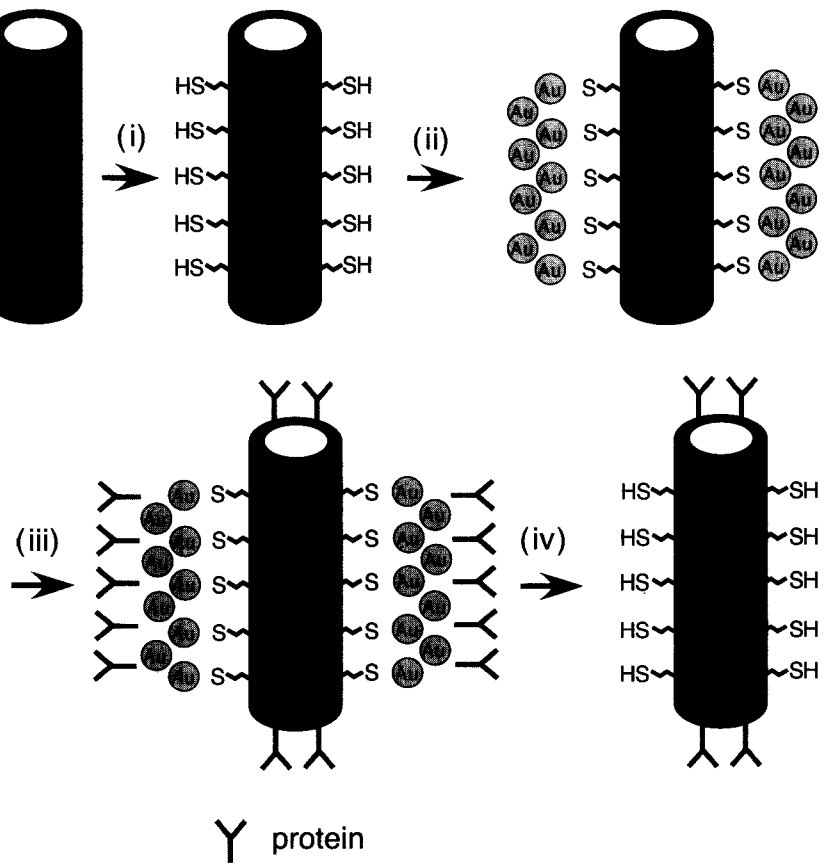


Fig. 7 Procedure to immobilize proteins at the ends of nanotubes using Au nanocrystals as protective masks. (From Ref. [57].) (*View this art in color at www.dekker.com.*)

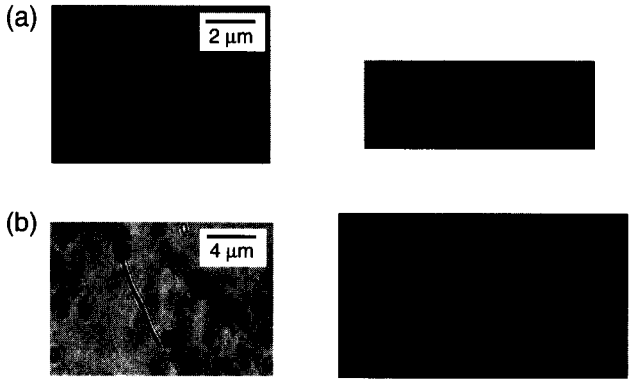


Fig. 8 Fluorescence micrographs of (a) the Au nanocrystal-masked nanotube incubated with sulforhodamine-labeled avidin [step (iii) in Fig. 7], (b) the nanotube with sulforhodamine-labeled avidin after etching the Au nanocrystals [step (iv) in Fig. 7]. (From Ref. [57].) (*View this art in color at www.dekker.com.*)

ends remained bound.[57] After sulforhodamine-labeled avidin was incubated into the Au nanocrystal-masked nanotube solution for 36 hr [step (iii) in Fig. 7], the protein was immobilized on the entire nanotube surface shown by fluorescence micrograph of the resulting nanotube (Fig. 8a). This fluorescence image was dramatically changed after the Au nanocrystals on the nanotube sidewall were etched by iodine in step (iv). After etching the Au nanocrystals from the nanotube surface, fluorescing areas of the sulforhodamine-labeled avidin were limited to the ends of the nanotube and fluorescence at the nanotube sidewall was diminished, as shown in Fig. 8b. These fluorescence micrographs confirm that the protein remained bound to the peptide nanotube ends after the chemical Au nanocrystal etching. This chemical etching process did not denature the avidin on the nanotubes and the nanotube ends could still recognize the complementary biotin after the etching process.[57] This result is particularly important to this proposed research because denaturing proteins at the nanotube ends via the chemical etching process could prevent the use of this protein nanotube as building blocks for nanodevices due to the loss of the biological recognition function. Our preliminary result indicates that these end-functionalized nanotubes reduced the nanotube aggregation about five times on the biotin-SAMs compared to the nanotubes whose entire surfaces were coated by avidin.[57]

The location-specific assembly of protein nanotubes was also achieved via biological recognitions between inorganic surfaces and phages.[58] Phages and natural protein nanotubes with 6.6 nm in diameter and approximately 880 nm in length (Fig. 9A) can be modified to have specific peptide sequences on the surfaces to recognize and bind specific inorganic surface after multiple cycles of the peptide selection processes. To identify appropriate peptide sequences of phages to recognize specific inorganic surfaces, Whaley and coworkers used a combinatorial library of genetically engineered bacteriophage.[58] For example, this peptide combinatory approach identified proteins that specifically bind GaAs (100) and they were self-assembled on the patterned GaAs surfaces (Fig. 10).[59] The specificity of binding exerted the level over the recognition of crystal orientations.

Those genetically engineered bacteriophage–nanocrystal complexes can also be arranged into ordered films.[60] For example, when phages were engineered to recognize ZnS, they mineralized highly crystalline ZnS nanocrystals at the end of the tubules in $ZnCl_2/Na_2S$ solutions (Fig. 9B,C). Under the optimized experimental conditions, those phages with ZnS nanocrystals were spontaneously assembled into self-supporting hybrid

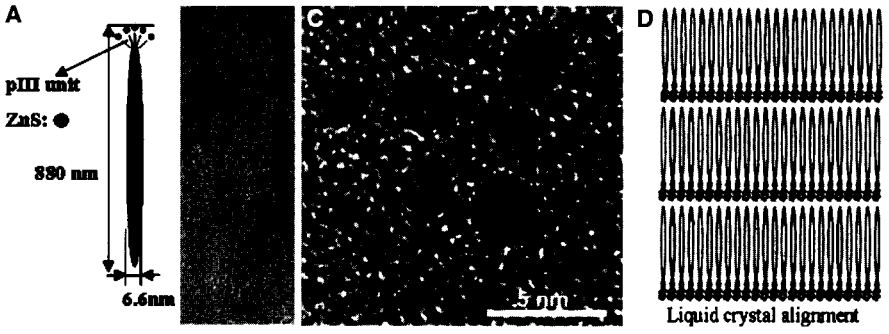


Fig. 9 (A) Illustration of phage and ZnS nanocrystals (the PIII peptide unit and the ZnS nanocrystal bound to the phage are not drawn in scale). (B) TEM image of an individual phage and ZnS nanocrystals. (C) HRTEM image of 0.01% phage–ZnS suspension, showing lattice fringe images of ZnS nanocrystals. (D) Illustration of liquid crystal alignment of the phage–ZnS nanocrystal complexes. (From Ref. [60]. Copyright (2002) American Association for the Advancement of Science.) (*View this art in color at www.dekker.com.*)

Fig. 10 Fluorescence image of tetramethyl rhodaimine-labeled phages selectively assembled on the patterned (100) GaAs surfaces. The red lines (1 μm across) correspond to GaAs and the black spaces (4 μm across) are SiO_2. (From Ref. [59]. Copyright (2002) National Academy of Science, U.S.A.) (*View this art in color at www.dekker.com.*)

films whose order structures were maintained over a centimeter-length scale (Fig. 9D). These films showed the liquid crystalline phase behaviors and they could be modulated by solvent concentrations and an external magnetic field.[60] The monodispersity of the genetically engineered phages and the ease in incorporating and aligning various types of inorganic nanocrystals by molecular cloning techniques and genetic selections are advantageous to apply the protein nanotube-based materials in the field of liquid crystals.

SENSING

As pointed out above, one of the key functions of proteins is to recognize specific molecules and ions. The molecular recognitions of proteins are dependent on the shapes of target molecules because multiple hydrogen bonds, electrostatic interactions, and hydrophobic interactions between target molecules and proteins in a well-defined spatial pattern provide the excellent selectivity of protein bindings.[61] Therefore it is natural to apply the molecular recognition of protein nanotubes to sense and diagnose target molecules.

A planar field effect transistor (FET) has been used mostly as a nanotube-based sensor configuration. Modification of the gate oxide in the carbon nanotube-based FET allowed monitoring a conductance change due to the adsorption of analytes.[62] The practical aspect of the FET is that the FET can serve as a highly sensitive real-time electrically based sensor. The sensitivity of carbon nanotube-based FET sensors was in the regime of single-molecule detection.[62]

Cui and coworkers assembled antigen nanowires, antigen-coated silicon nanowires, as building blocks into the FET sensors.[63] The single-crystal boron-doped (p-type) silicon nanowires, prepared by a nanocluster-mediated vapor–liquid–solid growth method,[64] were aligned by a microfluidic technique between two electrodes, a source and a drain (Fig. 11).[29] The reversible, real-time antigen-nanowire sensor was fabricated by functionalizing the silicon nanowire template with biotin via the drop deposition, and this sensor system could detect conductivity changes under m-antibiotin adsorption on the functionalized nanowires.[63] A linear change in the nanowire conductance as a function of m-antibiotin concentration was observed as shown in Fig. 11. The correlation between linear current (I) vs. voltage (V) was transformed to the conductance (dI/dV). The detection of m-antibiotin in this sensor system was achieved below 10 nM, while the detection limit for streptavidin could reach in the order of pM.[63]

MECHANICAL MOTION

There have been major attempts to synthesize chemically powered nanoscale engines driven by light or chemical energy.[65,66] The weaknesses of the synthetic nanoscale motors are as follows: 1) they often stall when operated under load; and 2) they have a low cycle time.[61] In contrast, motor proteins are better engines at this point as the man-made motors are fueled by electrical or thermal energy with an efficiency of around 30% while protein motors are driven by direct conversion from chemical energy with an efficiency as high as 90%.[67] Therefore motor proteins have been studied extensively not only to understand the mechanism for their mechanical

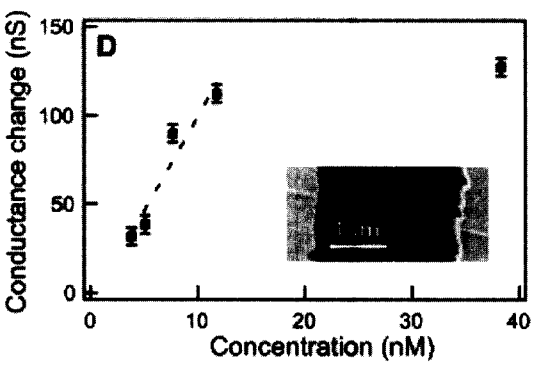


Fig. 11 Plot of the conductance change of a biotin-modified Si nanowire vs. m-antibiotin concentration. (Inset) SEM image of a typical FET configuration: a Si nanowire connecting two electrodes, a source and a drain. (From Ref. [63]. Copyright (2001) American Association for the Advancement of Science.)

motions regulating cell activities but also to fabricate high-performance nanoscale engines. There are two types of well-studied protein motors, translational motors and rotary motors. Translation protein motors transport on the complimentary protein nanotubes. While development of rotary protein-based nanodevices has also been very active,[68] rotary protein motors are not discussed in this chapter.

Conformational changes of motor proteins produce mechanical forces and motions fueled by the interconversion of chemical energy.[69] Translational motor proteins move along their complementary nanotubes via the energy conversion. Microtubules, protein nanotubes in diameter of 25 nm, are tracks to guide a translational motor protein, kinesin. Kinesins transport vesicles, chromosomes, and protein complexes using these tracks throughout the intracellular environment.[70] For example, kinesins transfer encapsulated neurotransmitters from cores of nerve cells to synaptic clefts in order to excite next nerve cells.[61]

Kinesins are elongated heterotetrameric proteins and contain adenosine triphosphate (ATP) and microtubule binding sites.[30] The side-by-side motor domains alternately bind and release the tubulin subunits of a microtubule, α- and β-tubulins, in a walking type of motion.[71] The walking motion is fueled by the hydrolysis of ATP from adenosine diphosphate (ADP), initiated by a proton gradient across a cell membrane. The kinetically more dynamic end of the microtubule is terminated with the β-tubulin (black circles in Fig. 12a) whose relative polarity is positive, whereas the opposite end is terminated with α-tubulin (white circles in Fig. 12a) whose relative polarity is negative. The overall structural polarity of microtubules is recognized by motor proteins, which determine the direction of their motion along the microtubule surface as kinesin steps only toward the plus end.[72] Myosins are another family of translational protein motors that move along actin nanotubes for cell motility and muscle contraction.[61]

While there have been extensive effort to manipulate this molecular transport machinery outside cells and fabricate as moving building blocks of nanodevices driven by kinesin/myosin motor proteins,[71] their applications require to overcome technical problems such as the orientation of force generation toward the desired direction, the optimization of surface conditions and geometries, and the stability of the motor protein–protein nanotube systems.[73]

Motor protein–protein nanotube integrated systems have been fabricated in two configurations: 1) moving motor proteins on protein nanotube tracks; and 2) moving protein nanotubes on motor protein-patterned surfaces.

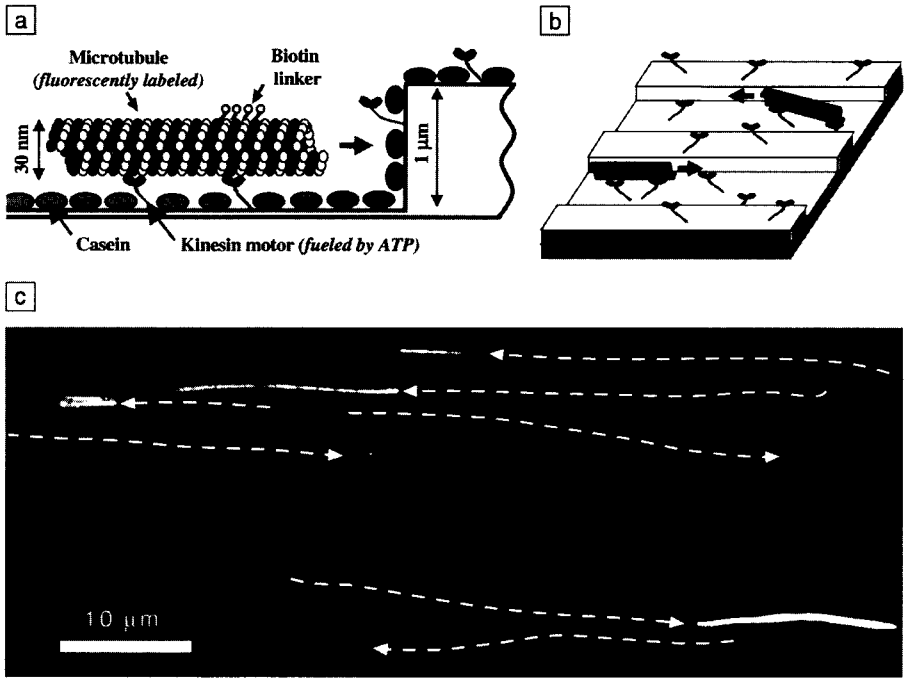


Fig. 12 (a) Illustration of the microtubule transport on kinesin motor surfaces. (b) Illustration of a typical geometry of the kinesin motor track. (c) Micrograph of photolabeled microtubules moving in channels (2 μm wide) on a patterned polyurethane substrate. The dotted lines are the paths of individual microtubules. (From Ref. [61]. Copyright (2002) the Materials Research Society.)

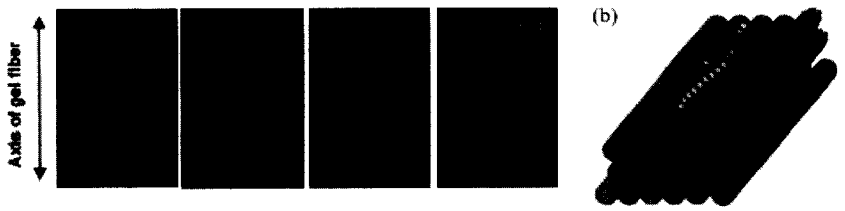


Fig. 13 (a) Sequential fluorescence images of the cross-linked actin gel movement on the oriented myosin gels. (b) Illustration of the actin gel movement on the myosin gels. (From Ref. [67]. Copyright (2002) Wiley-VCH Verlag GmbH.) (*View this art in color at www.dekker.com.*)

The first configuration is analogous to the geometry found in cells.[70] When kinesins and microtubules were integrated to develop nanodevices, the majority of effort has been focused on positioning and alignment of microtubules.[30,70,71,73] The microtubules must be aligned precisely in order to transport cargoes attached to motor proteins toward the proper destinations. For example, Limberis and Stewart transported silicon microchips with genetically engineered kinesin motors to particular locations guided by patterned microtubules on substrates.[71]

The second configuration is the inversed geometry used in nature. Various surface topologies with lithographically fabricated ridges and valleys were examined as motor protein tracks to guide protein nanotubes inside the ridges or valleys (Fig. 12b). The difficulty in this scheme is that microtubules tend to escape from the valleys by binding the sidewalls of valleys, slowly climb up, and then move toward different directions.[74] This escaping motion of microtubules on a kinesin-coated polyurethane surface patterned with 2-μm-wide and 1-μm-deep channels is shown in Fig. 12c. In this figure, the bottom of the channels is in focus and the plateaus in between are out of focus. Some microtubules were staying in the channels but most of them moved up and down in the grooves.[74] The patterned myosin surfaces were shown to direct the actin nanotube movement in one direction more effectively.[75] Recently, the myosin gel substrate also appeared to regulate the movement of the cross-linked actin nanofibers in one direction (Fig. 13).[67] The optimal solution to engineer motor protein/protein nanotube tracks seems to be an optimized combination of chemical and topological track patterns,[74] and the myosin/actin system may be chemically favorable to reduce the nanotube sliding motions.

Those motor protein/protein nanotube systems have potential to be applied in the fields of mechanical engineering, biomedical engineering, and electronics. Coupling motors/nanotubes with gears or rotors may be used as switches or valves in microfluidic channels.[70] It is encouraging that recently motion of microtubules can be turned on and off sequentially on patterned motor protein substrates triggered by UV-induced release of caged ATP via the enzymatic ATP degradation, which is a crucial feature to develop the switches and the valves.[74] Motor protein-based microgenerators, pumps, and biosensors can be applied in medical implants.[71] For example, the motor protein-based microgenerators can power implanted microdevices for monitoring physiological processes because the mechanochemical energy of motor proteins is transduced to an electrical current. In any applications described above, alignments and chemical controls of protein nanotubes will be crucial to develop nanodevices from these smart components.

CONCLUSION

Protein nanotubes have distinguished characters compared with other nanotubes. Obviously, protein-assembled nanotubes are biocompatible and broad biological applications can be expected. Molecular recognition is another unique function in protein nanotubes. The molecular recognition, such as an antibody–antigen recognition, images the complementarity of a lock and a key—the lock being antibody and the key being antigen that is recognized via the spatial fit and noncovalent interactions to form the specific antibody–antigen complex. This specificity will make protein nanotubes smart building blocks for high-performance chemical sensors. When multiple antibodies are patterned wisely on surfaces, resulting device configurations, after antigen-nanotubes recognize and position onto their complementary regions, will be achieved via their robust self-assembling processes. This approach may build certain device configurations in an economical and simple manner. Coupling between protein nanotubes and motor proteins certainly offers intriguing potential to build nanodevices powered by converting chemical energy into mechanical force. When physical properties of protein nanotubes are not suitable to certain applications, protein nanotubes can be functionalized to

add the desired physical properties. For example, if conductive nanotubes need to be placed at a certain position, protein nanotubes can be coated with conductive metals via biomineralization after protein nanotubes are positioned at the desired position via molecular recognitions. To date, fabrication of protein nanotubes and organization of protein nanotubes have been investigated, respectively. Applications of protein nanotubes as building blocks in real-world nanodevices will be accelerated once those two outcomes are integrated.

Of course, this field is still extremely young and there are many concerns and hurdles to overcome. For example, proteins typically lose their functions within days or weeks unless they are frozen.[61] They degrade by various and often competing mechanisms, including thermal and enzymatic degradation as well as precipitation. For the applications in the coating and the arrangement of nanotubes, proteins need to be active during the nanotube assembly, which is normally achieved in a short time because of robust self-assembly. Therefore the protein degradation may not be a crucial issue for these applications. But when protein nanotubes are applied to the devices necessary to transform one structure to the other repeatedly or to detect target molecules over and over again, proteins must be evolved to survive under harsh environmental conditions.[76,77]

Various difficulties in applying protein nanotubes as building blocks of nanodevices will be overcome if we keep learning from nature. There are many natural protein nanotubes/nanofibers whose excellent properties are known, but their chemical structures remain unsolved. For example, a tensile strength of spider dragline silk is comparable with steel and its elasticity is comparable with rubber.[78] This unique combination makes the spider dragline silk mechanically superior to any other man-made materials. But it was just recently that the protein sequences and structures of the spider dragline silk were understood.[78] Progress in understanding unsolved natural protein nanotubes and mimicking those nanotubes will make protein nanotubes more attractive to apply them as building blocks for various nanometer-scaled devices.

ACKNOWLEDGMENT

I would like to thank my co-workers for the past four years for their hard work in the development of protein nanotubes and their applications: S. Pan, G.E. Douberly, P. Porrata, Yung-fou Chen, Lingtao Yu, Ramin Djalali, and Ipsita Banerjee. I also would like to acknowledge the financial supports from the National Science Foundation CAREER Award (EIA-0133493), the National Science Foundation NER program (ECS-0103430), and the U.S. Department of Energy (DE-FG-02-01ER45935).

REFERENCES

1. Tans, S.J.; Devoret, M.H.; Dai, H.; Thess, A.; Smalley, R.E.; Geerligs, L.J.; Dekker, C. Individual single-wall carbon nanotubes as quantum wires. Nature **1997**, *386* (6624), 474–477.
2. Bockrath, M.; Cobden, D.H.; McEuen, P.L.; Chopra, N.G.; Zettl, A.; Thess, A.; Smalley, R.E. Single-electron transport in ropes of carbon nanotubes. Science **1997**, *275* (5308), 1922–1925.
3. Dai, H.; Kong, J.; Zhou, C.; Franklin, N.; Tombler, T.; Cassel, A.; Fan, S.; Chapline, M. Controlled chemical routes to nanotube architectures, physics, and devices. J. Phys. Chem., B **1999**, *103* (51), 11246–11255.
4. Dai, H.; Wong, E.W.; Lieber, C.M. Probing electrical transport in nanomaterials: Conductivity of individual carbon nanotubes. Science **1996**, *272* (5261), 523–526.
5. Shim, M.; Kam, N.W.S.; Chen, R.J.; Li, Y.M.; Dai, H.J. Functionalization of carbon nanotubes for biocompatibility and biomolecular recognition. Nano Lett. **2002**, *2* (4), 285–288.
6. Rueckes, T.; Kim, K.; Joselevich, E.; Tseng, G.Y.; Cheung, C.-L.; Lieber, C.M. Carbon nanotube-based nonvolatile random access memory for molecular computing. Science **2000**, *289* (5476), 94–97.
7. Diehl, M.R.; Yaliraki, S.N.; Beckman, R.A.; Barahona, M.; Heath, J.R. Self-assembled, deterministic carbon nanotube wiring networks. Angew. Chem., Int. Ed. Engl. **2001**, *41* (2), 353–356.
8. Bachtold, A.; Hadley, P.; Nakanishi, T.; Dekker, C. Logic circuits with carbon nanotube transistors. Science **2001**, *294* (5545), 1317–1320.
9. Collins, P.G.; Arnold, M.S.; Avouris, P. Engineering carbon nanotubes and nanotube circuits using electrical breakdown. Science **2001**, *292* (5517), 706–709.
10. Cheung, C.L.; Kurtz, A.; Park, H.; Lieber, C.M. Diameter-controlled synthesis of carbon nanotubes. J. Phys. Chem., B **2002**, *106* (10). 2429–2433.
11. Chattopadhyay, D.; Lastella, S.; Kim, S.; Papadimitrakopoulos, F. Length separation of Zwitterion-functionalized single wall carbon nanotubes by GPC. J. Am. Chem. Soc. **2002**, *124* (5), 728–729.
12. Doorn, S.K.; Fields, R.E.; Hu, H.; Hamon, M.A.; Haddon, R.C.; Selegue, J.P.; Majidi, V. High resolution capillary electrophoresis of carbon nanotubes. J. Am. Chem. Soc. **2002**, *124* (12), 3169–3174.
13. Zhao, B.; Hu, H.; Niyogi, S.; Itkis, M.E.; Hamon, M.A.; Bhowmik, P.; Meier, M.S.; Haddon, R.C. Chromatographic purification and properties of soluble single-walled carbon nanotubes. J. Am. Chem. Soc. **2001**, *123* (47), 11673–11677.

14. Joselevich, E.; Lieber, C.M. Vectorial growth of metallic and semiconducting single-wall carbon nanotubes. Nano Lett. **2002**, *2* (10), 1137–1141.
15. Bahr, J.L.; Tour, J.M. Covalent chemistry of single-wall carbon nanotubes. J. Mater. Chem. **2002**, *12* (7), 1952–1958.
16. Duan, X.; Huang, Y.; Cui, Y.; Wang, J.; Lieber, C.M. Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices. Nature **2001**, *409* (6816), 66–69.
17. Yu, J.-Y.; Chung, S.-W.; Heath, J.R. Silicon nanowires: Preparation, device fabrication, and transport properties. J. Phys. Chem., B **2000**, *104* (50), 11864–11870.
18. Mbindyo, J.K.N.; Mallouk, T.E.; Mattzela, J.B.; Kratochvilova, I.; Razavi, B.; Jackson, T.N.; Mayer, T.S. Template synthesis of metal nanowires containing monolayer molecular junctions. J. Am. Chem. Soc. **2002**, *124* (15), 4020–4026.
19. Shelimov, K.B.; Davydov, D.N.; Moskovits, M. Template-grown high-density nanocapacitor arrays. Appl. Phys. Lett. **2000**, *77* (11), 1722–1724.
20. Gudiksen, M.S.; Wang, J.F.; Lieber, C.M. Synthetic control of the diameter and length of single crystal semiconductor nanowires. J. Phys. Chem., B **2001**, *105* (19), 4062–4064.
21. Li, Y.M.; Kim, W.; Zhang, Y.G.; Rolandi, M.; Wang, D.W.; Dai, H.J. Growth of single-walled carbon nanotubes from discrete catalytic nanoparticles of various sizes. J. Phys. Chem., B **2001**, *105* (46), 11424–11431.
22. Huynh, W.U.; Dittmer, J.J.; Alivisatos, A.P. Hybrid nanorod-polymer solar cells. Science **2002**, *295* (5564), 2425–2427.
23. Murphy, C.J.; Jana, N.R. Controlling the aspect ratio of inorganic nanorods and nanowires. Adv. Mater. **2002**, *14* (1), 80–82.
24. Mayers, B.; Gates, B.; Yin, Y.D.; Xia, Y.N. Large-scale synthesis of monodisperse nanorods of Se/Te alloys through a homogeneous nucleation and solution growth process. Adv. Mater. **2001**, *13* (8), 1380–1384.
25. Landes, C.F.; Link, S.; Mohamed, M.B.; Nikoobakht, B.; El-Sayed, M.A. Some properties of spherical and rod-shaped semiconductor and metal nanocrystals. Pure Appl. Chem. **2002**, *74* (9), 1675–1692.
26. Smith, P.A.; Nordquist, C.D.; Jackson, T.N.; Mayer, T.S.; Martin, B.R.; Mbindyo, J.; Mallouk, T.E. Electric field assisted assembly and alignment of metallic nanowires. Appl. Phys. Lett. **2000**, *77* (9), 1399–1401.
27. Ural, A.; Li, Y.M.; Dai, H.J. Electric field-aligned growth of single-walled carbon nanotubes on surfaces. Appl. Phys. Lett. **2002**, *81* (18), 3464–3466.
28. Hone, J.; Llaguno, M.C.; Nemes, N.M.; Johnson, A.T.; Fischer, J.E.; Walters, D.A.; Casavant, M.J.; Schmidt, J.; Smalley, R.E. Electrical and thermal transport properties of magnetically aligned single wall carbon nanotube films. Appl. Phys. Lett. **2000**, *77* (5), 666–668.
29. Huang, Y.; Duan, X.F.; Wei, Q.Q.; Lieber, C.M. Directed assembly of one-dimensional nanostructures into functional networks. Science **2001**, *291* (5504), 630–633.
30. Limberis, L.; Magda, J.J.; Stewart, R.J. Polarized alignment and surface immobilization of microtubules of kinesin-powered nanodevices. Nano Lett. **2001**, *1* (5), 277–280.
31. Liu, J.; Casavant, M.J.; Cox, M.; Walters, D.A.; Boul, P.; Lu, W.; Rimberg, A.J.; Smith, K.A.; Colbert, D.T.; Smalley, R.E. Controlled deposition of individual single-walled carbon nanotubes on chemically functionalized templates. Chem. Phys. Lett. **1999**, *303* (1–2), 125–129.
32. Krupke, R.; Malik, S.; Weber, H.B.; Hampe, O.; Kappes, M.M.; v. Lohneysen, H. Patterning and visualizing self-assembled monolayers with low-energy electrons. Nano Lett. **2002**, *2* (10), 1161–1164.
33. Almassy, R.J.; Janson, C.A.; Hamlin, R.; Xuong, N.H.; Eisenberg, D. Novel subunit–subunit interactions in the structure of glutamine-synthetase. Nature **1986**, *323* (6086), 304–309.
34. Douberly, G.E.J.; Pan, S.; Walters, D.; Matsui, H. Fabrication of protein tubules: Immobilization of proteins on peptide tubules. J. Phys. Chem., B **2001**, *105* (9), 7612–7618.
35. Huang, W.; Taylor, S.; Fu, K.; Lin, Y.; Zhang, D.; Hanks, T.W.; Rao, A.M.; Sun, Y.-P. Attaching proteins to carbon nanotubes via diimide-activated amidation. Nano Lett. **2002**, *2* (4), 311–314.
36. Zhang, S.; Marini, D.M.; Hwang, W.; Santoso, S. Design of nanostructured biological materials through self-assembly of peptides and proteins. Curr. Opin. Chem. Biol. **2002**, *6* (6), 865–871.
37. Azamian, B.R.; Davis, J.J.; Coleman, K.S.; Bagshaw, C.B.; Green, M.L.H. Bioelectrochemical single-walled carbon nanotubes. J. Am. Chem. Soc. **2002**, *124* (43), 12664–12665.
38. Behrens, S.; Rahn, K.; Habicht, W.; Bohm, K.-J.; Rosner, H.; Dinjus, E.; Unger, E. Nanoscale particle arrays induced by highly ordered protein assemblies. Adv. Mater. **2002**, *14* (22), 1621–1625.
39. Dabrowski, M.D.; Chen, J.P.; Shi, H.Q.; Chen, W.C.; Atkins, W.M. Strategies for protein-based nanofabrication: Ni^{2+}-NTA as a chemical mask to control biologically imposed symmetry. Chem. Biol. **1998**, *5* (12), 689–697.

40. Schurke, P.; Freeman, J.C.; Dabrowski, M.J.; Atkins, W.M. Metal-dependent self-assembly of protein tubes from *Escherichia coli* glutamine synthetase. J. Biol. Chem. **1999**, *274* (39), 27963–27968.
41. Mishra, B.; Thomas, B.N. Phospholipid/protein cones. J. Am. Chem. Soc. **2002**, *124* (24), 6866–6871.
42. Whitesides, G.M. Self-assembling materials. Sci. Am. **1995**, *273* (3), 146–149.
43. Shimizu, T.; Kogiso, M.; Masuda, M. Vesicle assembly in microtubes. Nature **1996**, *383* (6600), 487–488.
44. Fenniri, H.; Deng, B.-L.; Ribbe, A.E.; Hallenga, K.; Jacob, J.; Thiyagarajan, P. Entropically driven self-assembly of multichannel rosette nanotubes. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99*, 6487.
45. Shimizu, T. Bottom-up synthesis and structural properties of self-assembled high-axial-ratio nanostructures. Macromol. Rapid Commun. **2002**, *23* (5–6), 311–331.
46. Demir, M.; Stowell, M.H.B. A chemoselective biomolecular template for assembling diverse nanotublar materials. Nanotechnology **2002**, *13* (4), 541–544.
47. Whitling, J.M.; Spreitzer, G.; Wright, D.W. A combinatorial and informatics approach to CdS nanoclusters. Adv. Mater. **2000**, *12* (18), 1377–1380.
48. Slocik, J.M.; Moore, J.T.; Wright, D.W. Monoclonal antibody recognition of histidine-rich peptide encapsulated nanoclusters. Nano Lett. **2002**, *2* (3), 169–173.
49. Shenton, W.; Douglas, T.; Young, M.; Stubbs, G.; Mann, S. Inorganic–organic nanotube composites from template mineralization of tobacco mosaic virus. Adv. Mater. **1999**, *11* (3), 253–256.
50. Douglas, T.; Young, M. Virus particles as templates for materials synthesis. Adv. Mater. **1999**, *11* (8), 679–681.
51. Field, M.; Smith, C.J.; Awshalom, D.D.; Mayes, E.L.; Davis, S.A.; Mann, S. Ordering nano-meter scale magnets using bacterial thread templates. Appl. Phys. Lett. **1998**, *73* (12), 1739–1741.
52. Djalali, R.; Chen, Y.-F.; Matsui, H. Au nanowire fabrication from sequenced histidine-rich peptide. J. Am. Chem. Soc. **2002**, *124* (46), 13660–13661.
53. Djalali, R.; Chen, Y.-F.; Matsui, H. Au nanocrystal growth on nanotubes controlled by conformations and charges of sequenced peptide templates. J. Am. Chem. Soc. **2003**, *125* (19), 5873–5879.
54. Beverly, K.C.; Sample, J.L.; Sampaio, J.F.; Remacle, F.; Heath, J.R.; Levine, R.D. Quantum dot artificial solids: Understanding the static and dynamic role of size and packing disorder. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99*, 6456.
55. Matsui, H.; Dagani, R. Building from the bottom up. C&EN **October 16, 2000**, 27.
56. Matsui, H.; Porrata, P.; Douberly, G.E.J. Protein tubule immobilization on self-assembled monolayers on Au. Nano Lett. **2001**, *1* (9), 461–464.
57. Banerjee, I.A.; Yu, L.; Matsui, H. Location-specific biological functionalization on nanotubes: Attachment of proteins at the ends of nanotubes using Au nanocrystal masks. Nano Lett. **2003**, *3* (3), 283–287.
58. Whaley, S.R.; English, D.S.; Hu, E.L.; Barbara, P.F.; Belcher, A.M. Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly. Nature **2000**, *405*, 665.
59. Seeman, N.C.; Belcher, A.M. Emulating biology: Building nanostructures from the bottom up. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99* (2), 6451–6455.
60. Lee, S.W.; Mao, C.B.; Flynn, C.E.; Belcher, A.M. Ordering of quantum dots using genetically engineered viruses. Science **2002**, *296* (5569), 892–895.
61. Vogel, V. Learning from proteins how to enhance the performance of synthetic nanosystems. MRS Bull. **2002**, *27* (12), 972–978.
62. Kong, J.; Franklin, N.R.; Zhou, C.W.; Chapline, M.G.; Peng, S.; Cho, K.J.; Dai, H.J. Nanotube molecular wires as chemical sensors. Science **2000**, *287* (5453), 622–625.
63. Cui, Y.; Wei, Q.Q.; Park, H.K.; Lieber, C.M. Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. Science **2001**, *293* (5533), 1289–1292.
64. Cui, Y.; Duan, X.; Hu, J.; Lieber, C.M. Doping and electrical transport in silicon nanowires. J. Phys. Chem., B **2000**, *104* (22), 5213–5216.
65. Feringa, B.L. In control of motion: From molecular switches to molecular motors. Acc. Chem. Res. **2001**, *34* (6), 504–513.
66. Stoddart, J.F. Molecular machines. Acc. Chem. Res. **2001**, *34* (6), 410–411.
67. Kakugo, A.; Sugimoto, S.; Gong, J.P.; Osada, Y. Gel machines constructed from chemically cross-linked actins and myosins. Adv. Mater. **2002**, *14* (16), 1124–1126.
68. Soong, R.K.; Bachand, G.D.; Neves, H.P.; Olkhovets, A.G.; Craighead, H.G.; Montemagno, C.D. Powering an inorganic nanodevice with a biomolecular motor. Science **2000**, *290* (5496), 1555–1558.
69. Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton*; Sinauer Associates, Publishers: Sunderland, MA, 2001.
70. Brown, T.B.; Hancock, W.O. A polarized microtubule array for kinesin-powered nanoscale assembly

and force generation. Nano Lett. **2002**, *2* (10), 1131–1135.
71. Limberis, L.; Stewart, R.J. Toward kinesin-powered microdevices. Nanotechnology **2000**, *11* (2), 47–51.
72. Schnapp, B.J.; Reese, T.S.; Nechttold, R. Kinesin is bound with high affinity to squid axon organelles that move to the plus-end of microtubules. J. Cell. Biol. **1992**, *119*, 389–399.
73. Stracke, R.; Bohm, K.J.; Burgold, J.; Schacht, H.-J.; Unger, E. Physical and technical parameters determining the functioning of a kinesin-based cell-free motor system. Nanotechnology **2000**, *11* (2), 52–56.
74. Hess, H.; Clemmens, J.; Qin, D.; Howard, J.; Vogel, V. Light-controlled molecular shuttles made from motor proteins carrying cargo on engineered surfaces. Nano Lett. **2001**, *1* (5), 235–239.
75. Nicolau, D.V.; Taguchi, T.; Taniguchi, H.; Yoshikawa, S. Negative and positive tone protein patterning on e-beam/deep-UV resists. Langmuir **1999**, *15* (11), 3845–3851.
76. Arnold, F.H. Combinatorial and computational challenges for biocatalyst design. Nature **2001**, *409* (6817), 253–257.
77. Schultz, P.G.; Lerner, R.A. Antibody catalysis: Completing the circle. Nature **2002**, *418*, 485.
78. Oroudjev, E.; Soares, J.; Arcidiacono, S.; Thompson, J.B.; Fossey, S.A.; Hansma, H.G. Segmented nanofibers of spider dragline silk: Atomic force microscopy and single-molecule force spectroscopy. Proc. Natl. Acad. Sci. U. S. A. **2002**, *99* (2), 6460–6465.

P

Proteins: Structure and Interaction Patterns to Solid Surfaces

P

Thomas J. Webster
Purdue University, West Lafayette, Indiana, U.S.A.

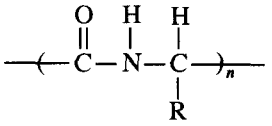
INTRODUCTION

Advances in several critical research fields (processing, catalytic, optical, actuation, electrical, mechanical fields, etc.) have started to benefit from new technological advancements in the area of nanotechnology.[1–22] Nanotechnology can be broadly defined as the use of materials and systems whose structures and components exhibit novel and significantly changed properties when control is gained at the atomic, molecular, and supramolecular levels. Specifically, such advances have been found for materials when particulate size is decreased to below 100 nm. However, to date, relatively few advantages have been described for biological applications (specifically, those involving protein interactions). This chapter elucidates several promising examples of the control nanophase materials have in manipulating interactions with proteins to improve traditional biological applications (such as for improving implantable devices, filtration systems, bioMEMS, etc.). Such advances were previously unimaginable with conventional materials possessing large micron size particulates. To describe the promise nanophase materials have in revolutionizing interactions with biology, this chapter will start with a review of protein structure and interaction with solid surfaces. The chapter will then continue with reports from the literature providing evidence of unique interactions proteins can have with nanophase materials.

PROTEIN STRUCTURE

Proteins are essential to health and, in fact, mean "of first importance" in Greek. This is for a very good reason because initial protein interactions with material surfaces control subsequent cell functions essential for tissue engineering/implant applications, antibody interactions for immunoassays and/or biodetection devices, and bacteria interactions involved in problems such as marine fouling and infection. Clearly, understanding protein size and structure will aid investigators in either maximizing or minimizing protein interactions with surfaces to improve these technologies.

Proteins are polyamides formed by a step reaction polymerization between the amino and carboxyl groups of amino acids:



where R is a characteristic side group and n is the number of repeating units. Depending on the side group, the molecular structure of the protein will drastically change. The simplest side group in protein structure is hydrogen, which will form the amino acid glycine (abbreviated Gly). The geometry of Gly is shown in Fig. 1 as a hypothetical flat sheet with a repeating distance of 0.72 nm.[23] Although the R side group in proteins can be diverse (as will soon be discussed), there are many commonalities between proteins, namely, size and structure in the nanometer regime. To date, the fundamental nanometer size of proteins has been largely ignored and not appreciated when considering approaches to control initial protein interactions with surfaces. Thus it is first important to describe protein structure to discuss how nanotechnology can be utilized to influence protein interactions. There are four important levels of protein structure to consider: primary, secondary, tertiary, and quaternary; these will be discussed in the sections that follow.

Primary Protein Structure

The primary structure of a protein is its linear sequence of amino acids. Each amino acid is linked to another through peptide bonds (Fig. 2).[23] Because there are 20 amino acids found in man (11 of which are synthesized in the human body: Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, Pro, Ser, and Tyr; and 9 of which are synthesized elsewhere but are essential to human diet: His, Ile, Leu, Lys, Met, Phe, Thr, Trp, and Val) (Fig. 3),[24] many sequences are possible to create numerous proteins with diverse properties. Moreover, these 20 amino acids are themselves diverse (Table 1)[25–27] and may be classified according to a number of categories: such as either aliphatic amino acids, amino acids with hydroxyl-containing or sulfur-containing side chains, aromatic amino acids, basic amino

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009098

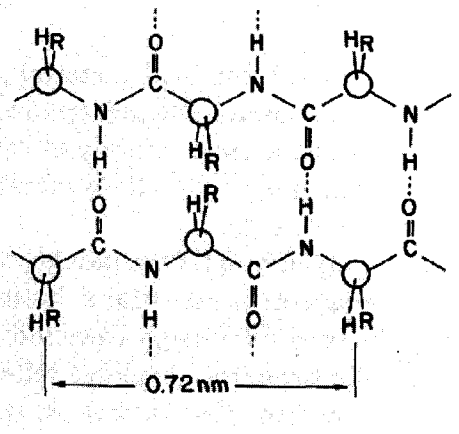


Fig. 1 The geometry of proteins as a hypothetical flat sheet with a repeating distance of 0.72 nm. (From Ref. [23].)

acids, cyclic amino acids, or acidic amino acids and their amides, as presented in Fig. 3. Other amino acid classifications that are often utilized when considering initial protein interactions with material surfaces are either nonpolar, polar, or ionized, labeled as either N, P, or I in Fig. 3, respectively.

Some amino acids have side chains (or residues) that carry no charge at any pH yet exhibit strong polar character (e.g., Ser and Thr).[23] Ionizable side chains vary from fairly acidic ones (such as Asp and Glu, which are fully negative at the physiological pH of 7.4) to more basic amino acids (such as the imidazole group in His, which has a positive charge at a pH of 7.4) and the still more basic amino acids that carry full charges at a pH of 7.4 (specifically, Lys and Arg).[23] Yet another group of amino acids has no acid, base, or polar character in their residues as shown by their generally much lower solubility in water (such as Ala, Val, Leu, Ile, Met, Trp, and Pro). For these reasons, it should not be surprising why proteins exist with a wide range of properties, as shown in Table 2, and this is just from considering protein primary structure.[25,26,28–32]

Secondary Structure

The secondary structure of proteins consists of regions of ordered structures in the protein chain (Fig. 4a and b).[23] Two main secondary structures of proteins are very common: the α-helix (Fig. 4a) and the β-pleated sheet (Fig. 4b). The degree of secondary structures in proteins varies to a large amount. For example, in structural proteins such as silk and wool, secondary structures are extensive and thus determine the overall shape of such proteins.[23] In addition, myoglobin and hemoglobin (proteins involved in oxygen transport and storage) have extensive α-helical secondary structure. In contrast, some proteins possess very little secondary structure such as the digestive enzyme chymotrypsin.

Both the α-helix and β-pleated sheet secondary protein structures are controlled by hydrogen bonding mechanisms, as illustrated in Fig. 4. Hydrogen bonding mechanisms are electrostatic attractions between oxygen of one chemical group and hydrogen of another chemical group. The α-helix results from the coiling of the protein chain such that the peptide bonds making up the backbone are able to form hydrogen bonds with each other (presented as curved lines in Fig. 4a). These hydrogen bonds are directed along the axis of the helix, and thus amino acid residues (portions of the amino acid not participating in the peptide bond) extend at right angles from the helix with minimal steric hindrance that further stabilizes the α-helix. The β-pleated sheet is a layering of protein chains in which each layer is held together by hydrogen bonds between the peptide links (presented as curved lines in Fig. 4b). The amino acid residues are situated at right angles to the sheets with, again, minimal steric hindrance that further stabilizes the β-pleated sheet.

Tertiary Structure

The tertiary structure of proteins is the overall three-dimensional shape of the protein (Fig. 5).[23] Protein tertiary structure can either possess little order for many structural proteins, or be extremely complicated for other proteins such as enzymes that fold up on themselves to form more complex structures. Clearly, the tertiary structure of proteins is a consequence of primary structure as it depends on spontaneous interactions between different amino acids and, under aqueous conditions, spontaneous interactions between amino acids and water as will be discussed. Because of the properties of the amino acids and the order in which amino acids are bonded (i.e., protein primary structure), select electrostatic interactions will be created for predictable tertiary structures. This automatic folding of proteins occurs even as proteins are

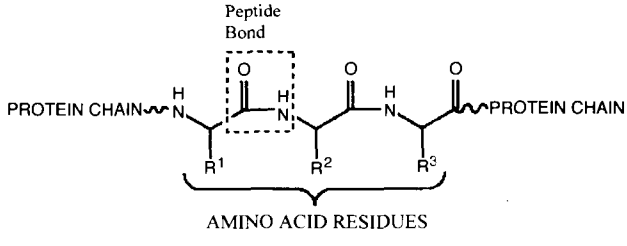


Fig. 2 Each amino acid is linked to another in proteins through peptide bonds. Note that the R group characteristic of each amino acid is often referred to as an amino acid residue. (From Ref. [23].)

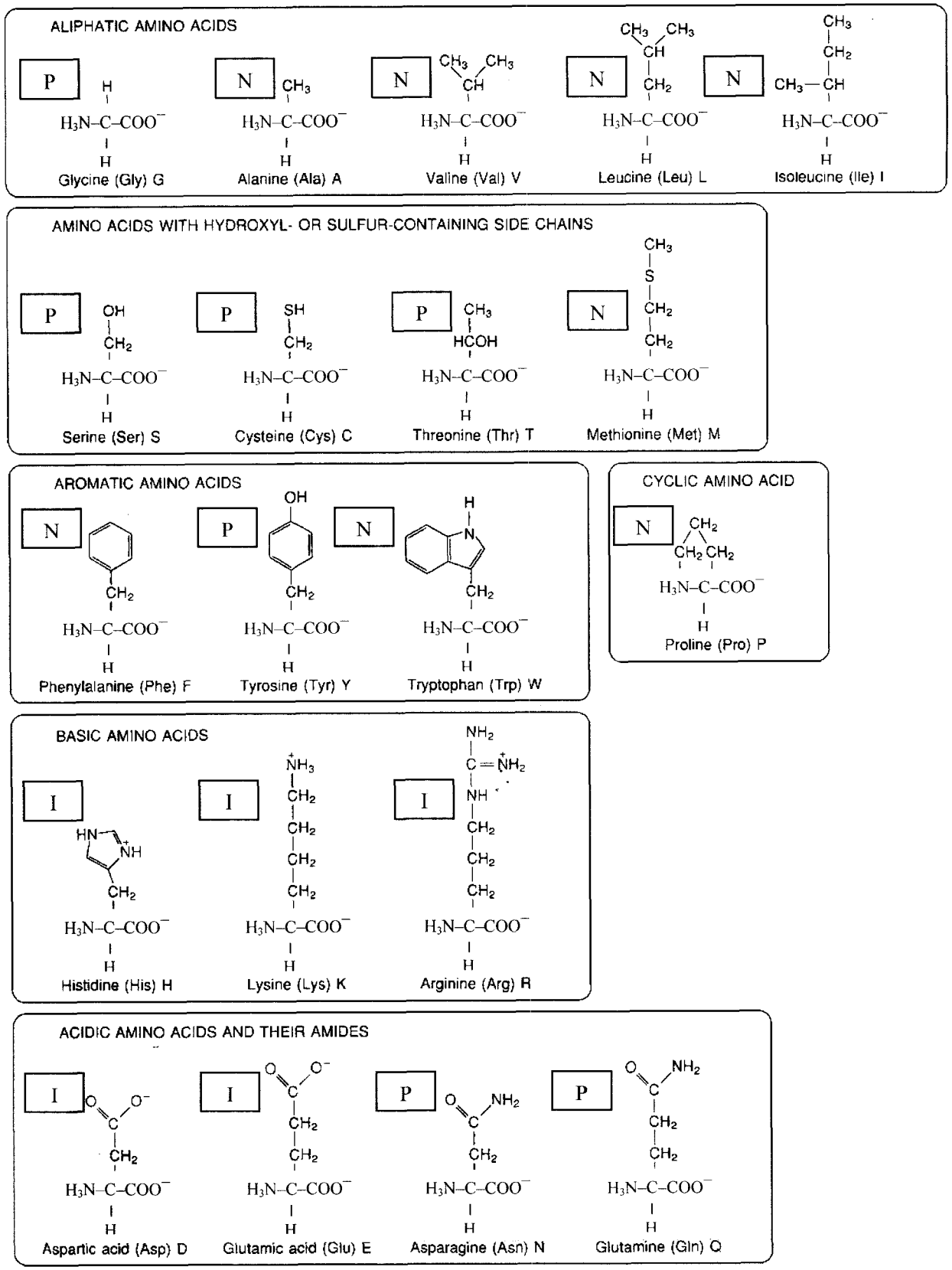


Fig. 3 Twenty common amino acids. N, P, and I are nonpolar, polar, and ionized residue categories, respectively, that appear to the left of each amino acid. (From Refs. [23] and [24].)

Table 1 Diverse properties of amino acid residues

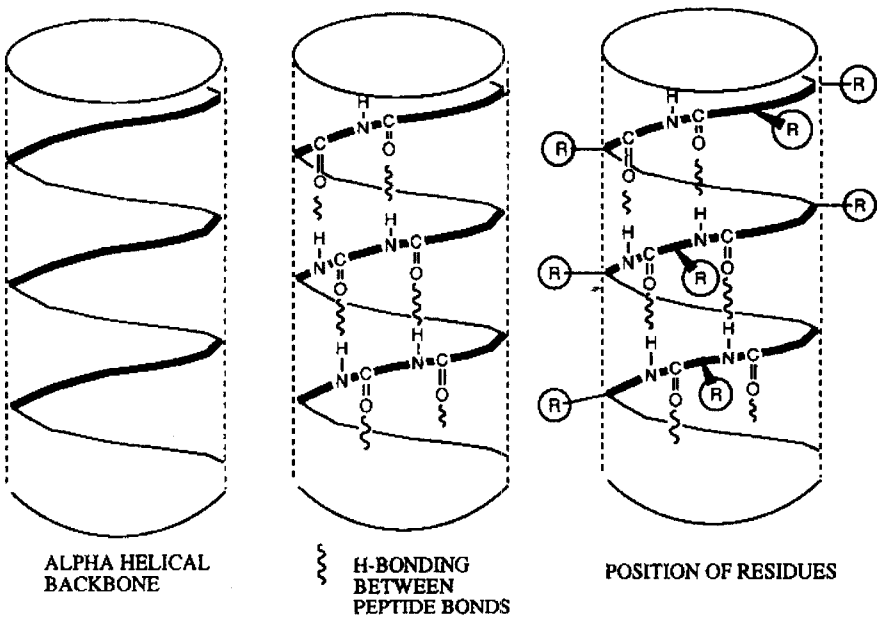
Amino acid (three-letter and one-letter abbreviations)	Charge	Hydrophobicity (kcal/mol)[a]	Surface tension (erg/cm^2/mol/l lowering of the surface tension of water)
Isoleucine (Ile or I)	Neutral	0.73	−15.2
Phenylalanine (Phe or F)	Neutral	0.61	−17.3
Valine (Val or V)	Neutral	0.54	−3.74
Leucine (Leu or L)	Neutral	0.53	−21.9
Tryptophan (Trp or W)	Neutral	0.37	−9.6
Methionine (Met or M)	Neutral	0.26	−3.01
Alanine (Ala or A)	Neutral	0.25	0.96
Glycine (Gly or G)	Neutral	0.16	1.12
Cysteine (Cys or C)	0 to −1	0.04	0.69
Tyrosine (Tyr or Y)	0 to −1	0.02	−15.1
Proline (Pro or P)	Neutral	−0.07	−0.49
Threonine (Thr or T)	Neutral	−0.18	0.59
Serine (Ser or S)	Neutral	−0.26	0.76
Histidine (His or H)	0 to 1	−0.40	1.03
Glutamic acid (Glu or E)	0 to −1	−0.62	0.86
Asparagine (Asn or N)	Neutral	−0.64	1.17
Glutamine (Gln or Q)	Neutral	−0.69	1.21
Aspartic acid (Asp or D)	0 to −1	−0.72	0.96
Lysine (Lys or K)	0 to 1	−1.1	0.92
Arginine (Arg or R)	0 to 1	−1.8	1.03

[a]Transfer from hydrophobic to more hydrophilic phase; increased positive value means more hydrophobic.
Source: Refs. [25–27].

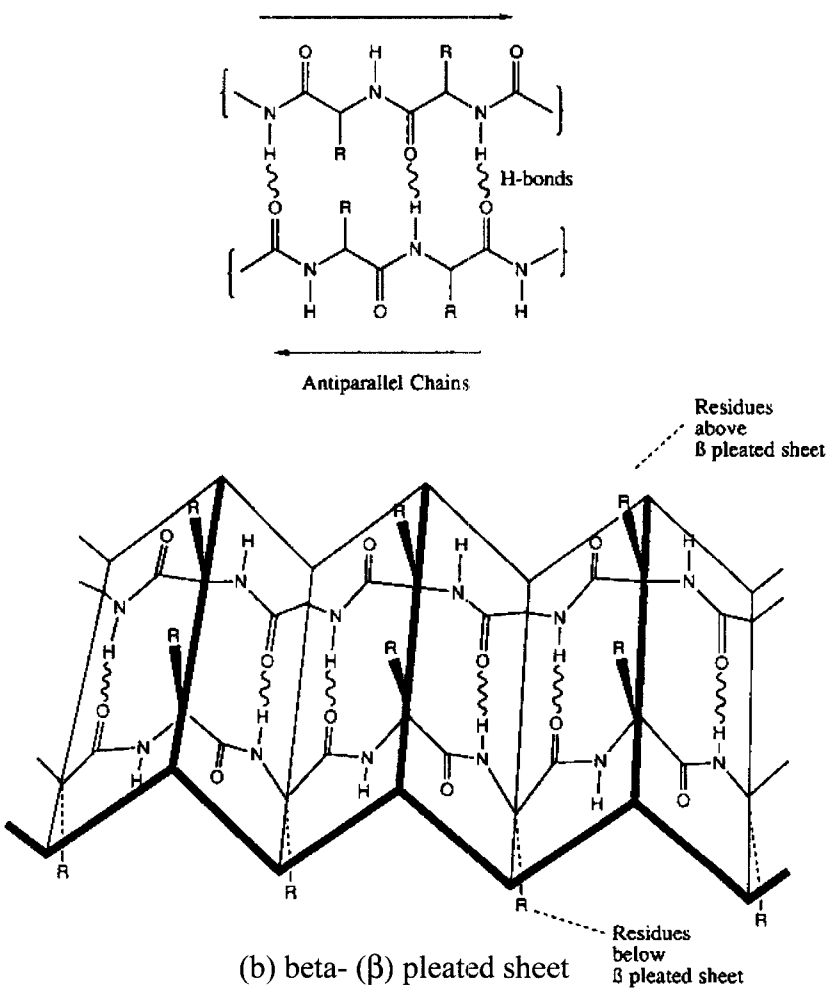
Table 2 Diverse properties of proteins

Protein	Function	Location	Size (kDa)	Shape (Å)	Stability	Surface activity
Albumin	Carrier	Blood	65	42 × 141	Denatures at 60°C	Low on polyethylene
Fibrinogen	Clotting	Blood	340	460 × 60 (trinodular string)	Denatures at 56°C	High on polyethylene
IgG	Antibody	Blood	165	T-shaped		Low on polyethylene
Lysozyme	Bacterial lysis	Tear; hen egg	14.6	45 × 30 (globular)	$\Delta G_n = -14$ kcal/mol	High on negatively charged surfaces
Hemoglobin	Oxygen carrier	Red blood cells	65	55 (spherical)	Normal form	Very high on polyethylene
Hemoglobin S	Oxygen carrier	Sickle red blood cells	65	55 (spherical)	Less than hemoglobin	Much higher air–water activity than hemoglobin
Myoglobin	Oxygen carrier	Muscle	16.7	45 × 35 × 25 (spherical)	$\Delta G_n = -12$ kcal/mol	
Collagen	Matrix factor	Tissue	285	3000 × 15 (triple helical rod)	Melts at 39°C	
Bacteriorhodopsin	Membrane protein		26	30–40 long		High at cell membrane
Tryptophan synthase α subunit (wild type)	Enzyme		27		$\Delta G_n = -8.8$ kcal/mol; denatures at 55°C	High air–water activity compared to ovalbumin
Tryptophan synthase variant α subunit	Enzyme		27		$\Delta G_n = -16.8$ kcal/mol	Much less active at air–water interface than wild type

Source: Refs. [25,26,28–32].



(a) alpha- (α) helix



(b) beta- (β) pleated sheet

Fig. 4 Secondary protein structures: (a) α-helix and (b) β-pleated sheet. (From Ref. [23].)

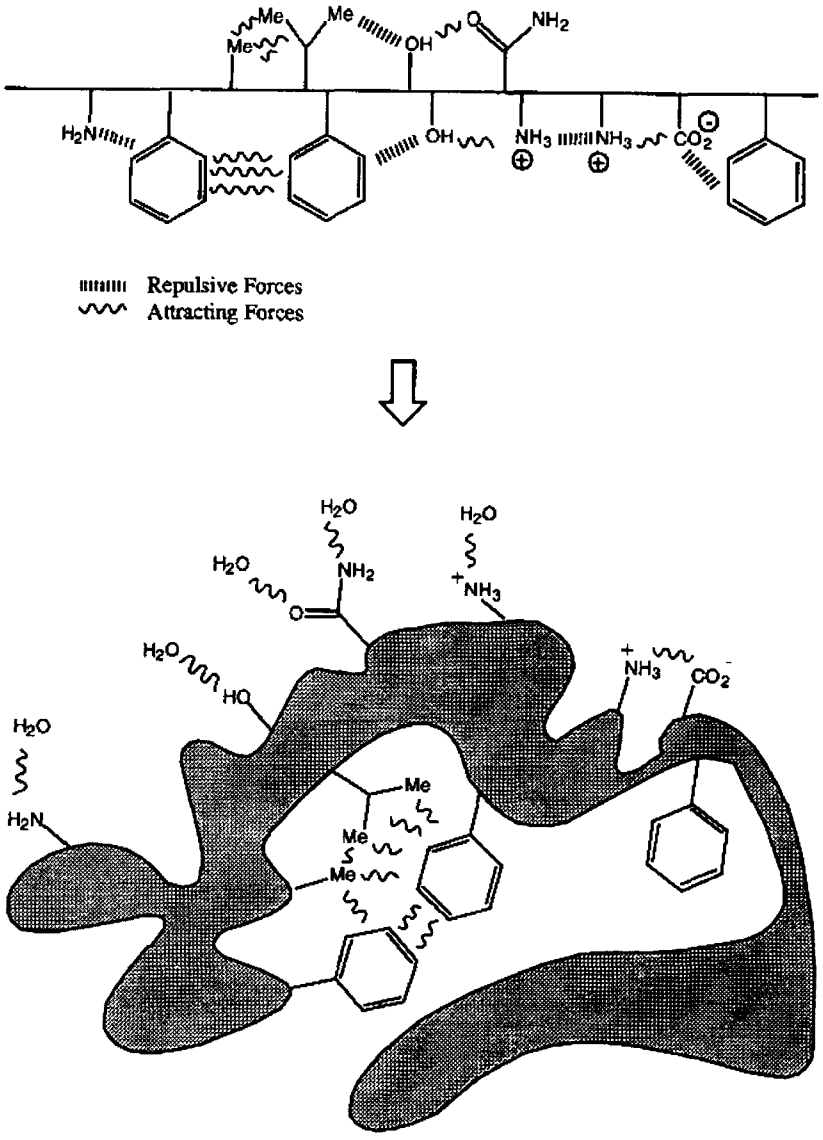


Fig. 5 Tertiary protein structure. (From Ref. [23].)

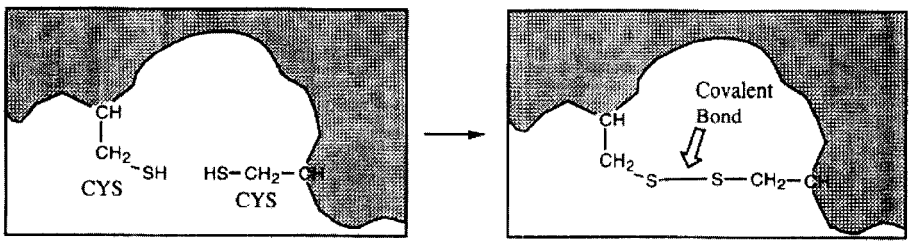
synthesized within ribosomes of cells. As individual proteins are synthesized, residues of each amino acid will either attract, repel, or remain indifferent to residues of another amino acid. Thus regions of the protein will twist and turn to minimize unfavorable and to maximize favorable interactions to form a consistent shape or three-dimensional tertiary structure, as illustrated in Fig. 5 (please note repulsive and attracting forces resulting in a favorable protein shape).

There are four main interactions among residues of amino acids that contribute to the tertiary structure of proteins, each with different strengths: covalent, ionic, hydrogen, and van der Waals bonds. Of these interactions, covalent bonds are the strongest, as indicated in Table 3.[23] For example, when two Cys amino acids come in close proximity, a covalent bond between sulfur in each residue can be formed as a result of oxidation (Fig. 6a).[23] This covalent bond can have a strength

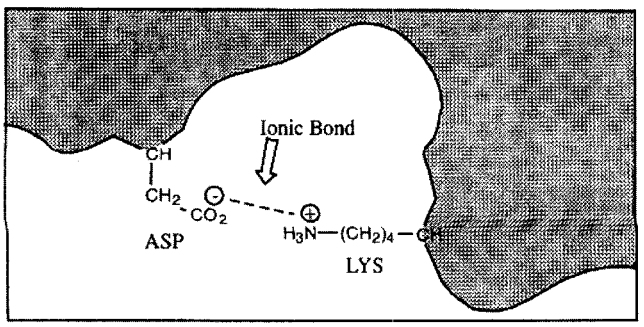
Table 3 Strengths of bonds controlling tertiary protein structure

Type of bond	Strength (kJ/mol)
Covalent (S—S)	250
Ionic	20
Hydrogen	7–40
Van der Waals	1.9

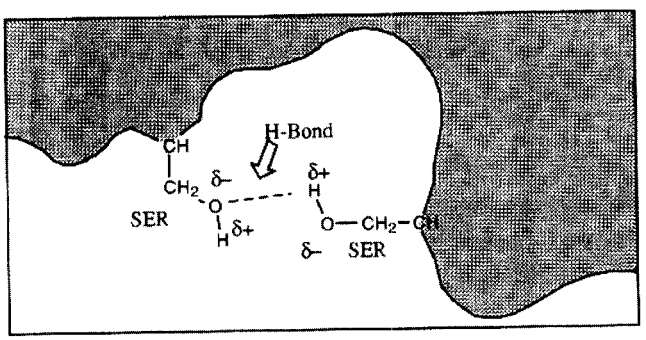
Source: Ref. [23].



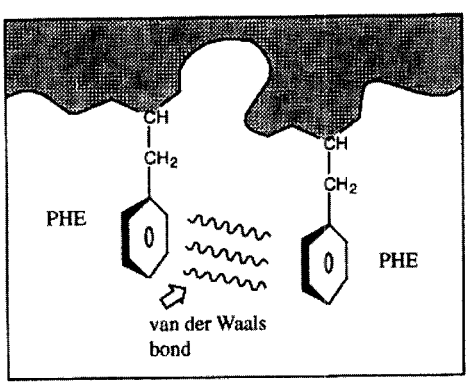
(a) covalent bonds



(b) ionic bonds



(c) hydrogen bonds



(d) van der Waals bond

Fig. 6 Four main interactions among residues of amino acids: (a) covalent, (b) ionic, (c) hydrogen, and (d) van der Waals bonds. (From Ref. [23].)

approaching 250 kJ/mol.[23] Ionic bonds are also a strong interaction force (approximately 20 kJ/mol),[23] which occurs between chemical groups having opposite charges. For example, ionic bonds are often formed between the carboxylate ion of an acidic residue (such as Asp) and the ammonium ion of a basic residue (such as Lys), as presented in Fig. 6b.[23] As previously mentioned, hydrogen bonds are formed between electronegative atoms, such as oxygen, and protons attached to electronegative atoms. Hydrogen bonding between two Ser amino acids is illustrated in Fig. 6c.[23] The forces can be quite strong in the range of 7–40 kJ/mol.[23] Lastly, van der Waals forces result from interactions between hydrophobic molecules (e.g., between two aromatic residues such as Phe, as presented in Fig. 6d). They result from unequal distributions of electrons between residues such that an area of high electron density on one residue has an attraction for an area of low electron density on another residue. These forces are the weakest of those discussed (1.9 kJ/mol).

Based on Table 3, one might expect that the most important or influential binding forces are those with the highest strength. On the contrary, generally speaking, the most important binding forces for the tertiary structure of proteins are those with the weakest strength (i.e., van der Waals and hydrogen bonding forces), whereas the least important forces are those with the highest strength (specifically, covalent and ionic bonding forces). One reason for this is because, compared with ionic and covalent bonding, there are many more opportunities for van der Waals and hydrogen bonding interactions in protein tertiary structure. A simple reexamination of amino acid chemical structures, as given in Fig. 3, can provide the answer as to why this is true. For example, the only covalent bond that is readily seen in protein tertiary structure is the disulfide bond.[23] The only amino acid capable of creating such a bond, when in juxtaposition to itself, is Cys.[23] In contrast, there are eight amino acids capable of promoting van der Waals interactions (specifically, Gly, Ala, Val, Leu, Ile, Phe, Pro, and Met).[23] However, it is important to note that for some proteins (small polypeptides such as the hormones vasopressin and oxytocin), covalent bonding mechanisms are more important, but for the majority of proteins, these bonds play a minor role in mediating tertiary structure compared with the other bonding mechanisms listed in Table 3. The same is true for ionic and hydrogen bonding mechanisms. Specifically, significantly more amino acids are capable of forming hydrogen bonds (eight: Ser, Thr, Cys, Asn, Gln, His, Tyr, and Trp) than ionic bonds (four: Asp, Glu, Lys, and Arg).[23]

Another reason for the increased importance of the weaker-strength interactions in controlling tertiary structure is the fact that proteins exist in aqueous media. Thus residues of amino acids must interact with water, which is a highly polar compound that forms strong hydrogen bonds (as presented in Fig. 5). In addition, water can also accept a proton to become positively charged and can form ionic bonds to several amino acids including Asp and Glu. This makes water an important medium for influencing the tertiary structure of proteins because it can form either hydrogen or ionic bonds with the following groups of amino acids: Ser, Thr, Cys, Asn, Gln, His, Tyr, Asp, Glu, Lys, and Arg.[23] These hydrophilic amino acids are either polar or ionizable, as given in Fig. 3. The remaining nonpolar amino acids (Ala, Val, Leu, Ile, Phe, Pro, and Met) are hydrophobic and thus are repelled by water. In this manner, it can be expected that the most stable tertiary structures of proteins in aqueous media are those where most of the hydrophobic and hydrophilic amino acids are on the inside and outside of the protein, respectively (Fig. 5). This is why most soluble proteins exhibit spherical or globular tertiary shapes, although fibrinogen (an important protein for blood clotting) does not follow this trend as it is more elongated than spherical.

Globular tertiary protein structures have several consequences in terms of amino acid interactions. The hydrophobic amino acids on the inside of the protein tertiary structure (avoiding interaction with water) are forced to interact with each other because of this tertiary structure. However, the number of ionic and hydrogen bonds contributing to the tertiary structure is reduced because the hydrophilic amino acids on the exterior of the protein form ionic and/or hydrogen bonds with water. For these reasons, hydrophobic—not hydrophilic—interactions control the tertiary structure of proteins. However, as will be discussed, hydrophilic and hydrophobic amino acid locations on the outer and inner cores of the protein, respectively, will drastically change as proteins initially come in contact with solid surfaces.

Of course, repulsion forces between amino acid residues in close proximity can also aid in protein tertiary structure. For example, residues that are hydrophilic (such as amino functional groups in Gln) would repel amino acid residues that are hydrophobic (such as an aromatic ring in Phe) when in close position. Similarly, two residues that are identical in charge will repel each other until a more favorable interaction between residues is created.

Quaternary Structure

Only proteins that possess numerous subunits have quaternary structure. How these subunits interact will determine the quaternary structure of that protein. An example of a well-studied quaternary protein structure is hemoglobin, which is composed of four protein subunits as described by secondary structure: two identical α subunits and two identical β subunits (Fig. 7).[24] Of

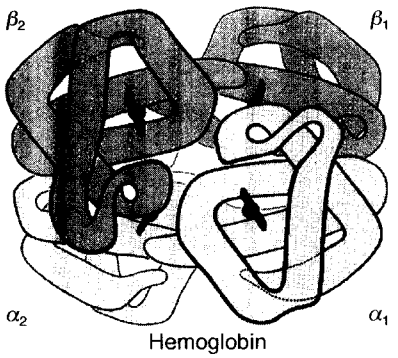


Fig. 7 Quaternary structure of hemoglobin. (From Ref. [24].)

course, interactions between amino acid residues on the exterior of the protein's tertiary structure will control quaternary structure. Thus ionic bonding mechanisms, which control the placement of hydrophilic amino acids on the exterior of protein tertiary structure in an aqueous environment, now play another important role in determining quaternary structure. Similarly, hydrophobic interactions have an essential role in determining quaternary structure. For example, it is clearly not possible for all hydrophobic amino acids to be placed on the interior of proteins; some may be present on the exterior. Thus a small hydrophobic area between adjacent subunits in the quaternary structure of proteins may be formed through hydrophobic interactions because these groups would have a preferred interaction of facing each other, rather than being exposed to the aqueous environment. This is another example of how important hydrophobic interactions are in controlling protein structure.

Although the structure of proteins has been well studied for a number of years, only in the past several decades have investigators begun to elucidate mechanisms of protein interactions with solid substrates. Knowledge of protein structure has certainly aided these researchers in understanding properties of surfaces that mediate protein interactions. "Protein Interactions with Surfaces" will describe some of these findings.

PROTEIN INTERACTIONS WITH SURFACES

Although proteins may interact in a wide range of media, such as at the air–water interface in the case of interfacial coagulation/foaming, this section will primarily discuss interactions at the solid–liquid interface because the most promising applications of nanophase materials in a biological context have been observed at this intersection.[33–42] Because proteins can interact with surfaces in two main ways—adsorption and orientation (or conformation)—the next sections will focus on these aspects beginning with adsorption.

Principles of Protein Adsorption

A great deal of attention has been focused on protein structure in aqueous media in this chapter. This is because soluble proteins present in biological fluids (e.g., blood plasma) are the type of proteins that are involved in immediate adsorption to surfaces.[43] In contrast, insoluble proteins that comprise tissues (such as collagen and elastin) are not normally free to diffuse to a solid surface; however, these proteins may appear on solid surfaces of implantable devices because of synthesis and deposition by cells.[25]

In as short a time that can be measured (i.e., less than 1 sec), soluble proteins become adsorbed to surfaces.[25] This is true for applications incorporating implants (such as orthopedic, vascular, etc.), bioseparation devices, immunoassays, catheters, and biosensors, or for any device involved in protein contact from a liquid surface to a solid surface. In seconds to minutes, a monolayer of adsorbed protein will form on solid surfaces.[25] Typical values for protein adsorption on solid surfaces are in the range of 1 $\mu g/cm^2$ and exhibit a plateau with respect to initial protein concentration.[25] That is, protein adsorption will reach a maximum and will not be influenced by higher initial bulk protein concentrations. Moreover, the surface concentration of proteins adsorbed on a material surface is often 1000 times more concentrated than in the bulk phase.[25] It is apparent that because there is a limited amount of space on a material surface, extreme competition exists for protein adsorption to solid surfaces. Depending on the two major driving forces for protein adsorption from liquid to solid surfaces (specifically, the relative bulk concentration of each protein in solution and the properties of the surface that control protein reactivity), the outcome of this competitive adsorption process is that the surface will be rich in some proteins while poor in others. Because proteins have vastly different properties and thus reaction to surfaces, as shown in Table 2,[25,26,28–32] protein adsorption will clearly be different depending on material chemistry, wettability, roughness, charge, etc.

It is also imperative to note that with some exceptions, protein adsorption to material surfaces is irreversible and thus leads to "immobilization" of specific proteins because they are, for the most part, not free to diffuse away.[44] Harsh treatments, such as using detergents (such as sodium dodecyl sulfate [SDS]), are usually required to remove adsorbed proteins from a material surface.[40]

Clearly, because of this efficiency, proteins have an inherent tendency to adsorb on surfaces as a tightly bound adsorbate, and removing unwanted protein adsorption (or fouling) from surfaces has consequently become an active area of research.[45–47]

Kinetics of Protein Adsorption

The kinetics of protein adsorption to solid surfaces can be described as a very fast initial phase that is diffusion-limited followed by a slower phase until steady state is realized.[25] A plot of the initial amount of protein adsorbed vs. the square root of time results in a linear characteristic typical of diffusion-controlled processes.[25] It can be speculated that during the later phase, protein adsorption is slowed because of the competition to locate a position sterically unhindered on the surface.[25]

Thermodynamics of Protein Adsorption

Because protein adsorption appears to be irreversible,[25,40,44] the thermodynamics of protein adsorption is difficult to characterize. However, there have been several studies providing insights into the thermodynamics of protein adsorption.[25,31,41] For example, some studies have determined direct measurements of the heat of adsorption for several proteins on a variety of surfaces, as indicated in Table 4.[25,31,41] Positive enthalpies have been documented on spontaneous adsorption, thus providing evidence that protein adsorption is entropically driven under certain circumstances. [Remembering that $\Delta G = \Delta H - T\Delta S$, with a positive ΔH (enthalpy), ΔS (entropy) must be positive to keep ΔG (free energy) negative to provide for a spontaneous reaction.] It is now generally agreed on that all protein adsorption processes are strongly driven by entropic changes.[25] The presence of entropic factors in protein adsorption to solid surfaces could arise because of the thermodynamics of water binding to the surface and the limited protein unfolding that occurs once adsorbed.[25]

Table 4 Positive enthalpy change of protein adsorption to surfaces

Protein	Surface	Enthalpy change of adsorption (mJ/m^2)
Horse serum albumin	α-Fe_2O_3	+1.9 at pH=5 +7.0 at pH=7
RNase	Negatively charged polystyrene with high surface charge density	+4 at pH=5 −2 at pH=11

Source: Refs. [25], [31], and [41].

Principles of Protein Orientation

As previously mentioned, protein adsorption is only one manner in which proteins interact with solid surfaces. The orientation of proteins in the adsorbed monolayer must also be considered because this interaction leads to extreme consequences in the ultimate function of a device.[36,44] For example, for implant, biosensor, or immunoassay applications, particular amino acids in proteins are well known to bind to a variety of cells, antibodies, or other agents that drastically influence device function. An example of the importance of protein orientation for the capture of cells is illustrated in Fig. 8.[48] Protein orientation will alter from surface to surface because proteins are not uniform in properties (Table 2) or structure. The existence of regions that are largely acidic/basic and hydrophobic/hydrophilic, and with select amino acids exposed to the aqueous media will greatly influence how that protein will adsorb to a surface, and thus its orientation, once adsorbed. As previously mentioned, proteins are not free to rotate once adsorbed because of multiple bonding mechanisms and thus immediately on adsorption, proteins are fixed in a preferred orientation or bioactivity.[25]

Under certain extreme conditions (e.g., conditions that are outside of the physiological range, or outside the range of 0–45°C and pH 5–8, and in aqueous solutions of about 0.15 M ionic strength), proteins may lose their normal structure.[25] In other words, under such conditions, the spherical or globular tertiary structure most soluble proteins assume in aqueous media will unfold or denature. The structure of denatured proteins has been described as a random coil structure similar to those found in synthetic polymers.[25] Because the structure of the protein has changed from that of a hydrophilic/hydrophobic exterior/interior to a more random arrangement, often times denatured proteins lose their solubility, become less dense (folded protein structures have densities of approximately 1.4 g/cm^3), and ultimately lose their bioactivity.[25] Although there have been many examples of protein denaturation in solution, in general, there have been few reports of full protein denaturation on material surfaces.[25] That is, generally, proteins adsorbed at the solid–liquid interface are not fully denatured and retain some degree of structure.

Properties That Influence Protein Interactions

Not only do properties of proteins determine the degree of protein interaction with surfaces, but properties of the

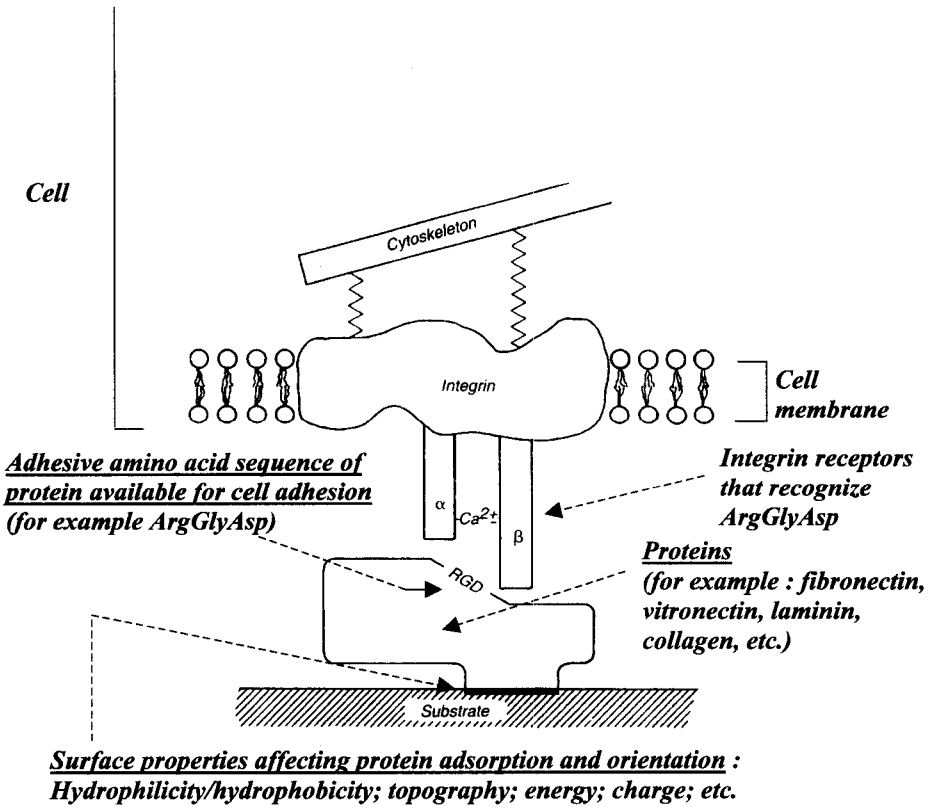


Fig. 8 Control of surface–protein interactions for capture of cells. (From Ref. [48].)

media and the surface (specifically, wettability, surface energy, chemistry, roughness, etc.) also influence the degree of protein interactions.[48] This is true because proteins are relatively large in size and have correspondingly great numbers of charged amino acid residues of varying acidity/basicity well distributed on the exterior of the protein. Therefore the inherent polyelectrolytic behavior of proteins becomes increasingly important depending on conditions of the aqueous media (such as alterations in pH and ionic strength). The polyelectrolytic property of proteins provides for exciting design criteria to engineer surfaces to either maximize or minimize protein interactions. Not surprisingly, at a neutral or slightly charged surface and at a pH in which the net charge on the protein is minimal (i.e., near the isoelectric pH), most proteins will exhibit maximum adsorption.[25] For surfaces with a large net charge, initial protein interactions will be dominated by the degree of the opposite charge on the surface.[25,48] Thus to develop a surface to maximize interactions with a protein displaying a large positive charge, one needs to design a surface with a large negative charge (such as one with numerous hydroxyl groups).

Simple consideration of the spatial organization of amino acids can also be used in the design of surfaces to enhance protein interactions.[39] As previously discussed, hydrophilic and hydrophobic amino acids are present primarily on the exterior and the interior of soluble proteins, respectively. This spatial arrangement has a direct consequence on the initial interaction of proteins with surfaces. For example, for the most part, a surface that initiates interactions with the exterior hydrophilic amino acid residues can promote adsorption. In contrast, for the interior hydrophobic amino acid residues to interact with a material surface, the soluble protein would have to unfold or lose tertiary structure, which is unlikely under normal conditions. For this reason, it can be reasoned that the types of amino acid residues that will initially react with a material surface are only those that exhibit largely polar properties (Fig. 3). Thus to minimize interactions with hydrophobic regions of proteins, one could design a surface with a large degree of hydrophobicity (such as polymers with numerous nonfunctionalized benzene rings).

Several studies have confirmed speculations that properties (such as chemistry, charge, topography, etc.) of surfaces dictate select interactions (specifically, the type, concentration, and conformation/bioactivity) of proteins.[43,49–53] It has been reported in the literature that changes in the type and concentration (up to 2100%, 74%, and 53% for albumin,[52] fibronectin,[54] and

vitronectin,[55,56] respectively) of protein adsorption on materials depend on surface properties such as chemistry (i.e., either polymer, metal, or ceramic), hydrophilicity/hydrophobicity, roughness, and surface energy. Specifically, maximum vitronectin,[56] fibronectin,[54] and albumin[43,52] adsorption was noted on hydrophilic ceramic surfaces with high surface roughness and/or energies. Despite these promising studies providing some level of understanding between material properties and protein adsorption, few (if any) studies have accurately correlated material properties to protein orientation. This lack of correlation may be because of a previous failure to create surfaces with features similar in size to those of proteins (i.e., at the nanometer level).

THE PROMISE OF NANOPHASE MATERIALS FOR ENHANCING PROTEIN INTERACTIONS

Nanophase materials are new formulations of materials that are composed of particulates of the same atoms but are fewer (less than tens of thousands) and smaller (less than 100 nm in diameter) than conventional forms (which contain several billions of atoms and have particulate sizes in microns to millimeters in diameter).[15] Representative images of nanophase compared with conventional materials are presented in Fig. 9.[35] The rationale, followed by experimental evidence that surface properties of nanophase materials enhance initial protein interactions, is expounded in the sections that follow.

Rationale

Surface properties (such as area, charge, and topography) depend on the particulate (such as grain) size of a material.[1,4,7,15,21,22] In this respect, nanophase materials that, by their very nature, possess higher surface areas with increased portions of surface defects (such as edge/corner sites) and/or grain boundaries[1,7] have special advantageous properties that remain largely unexplored for applications involving biology. To date, the increased surface reactivity of nanomaterials has been utilized for catalytic applications exclusively;[1,7] for example, compared with conventional (greater than 100 nm in average grain size) magnesium oxide (MgO), nanophase (i.e., 4 nm average grain size) MgO possessed an increased number of atoms at the surface, higher surface area (100–160 m^2/g compared with 200–500 m^2/g. respectively), less acidic OH^- groups (because of a much higher proportion of edge sites for the nanophase MgO to cause delocalization of electrons) (Fig. 10), increased adsorption of acidic species, and increased destructive adsorption of organophosphorous and chlorocarbons.[1,7] Because such promising results have been found when utilizing nanophase materials in catalytic applications, it is intriguing to ponder what promise these materials may have in applications involving protein interactions.

Experimental Evidence

Some of the best examples of how nanophase—compared with conventional—materials alter interactions with proteins are outlined in investigations of the potential use of nanostructured materials as the next generation of bone implants.[33–42] Reports in the literature have determined that new bone synthesis is enhanced on materials with grain sizes less than 100 nm.[33–42] Investigations of the underlying mechanisms revealed that the initial adsorbed concentration,[36] conformation,[40] and bioactivity[40] of proteins contained in blood serum were responsible for the select enhanced functions of osteoblasts (or bone-forming cells) on nanophase materials. Of first importance, the

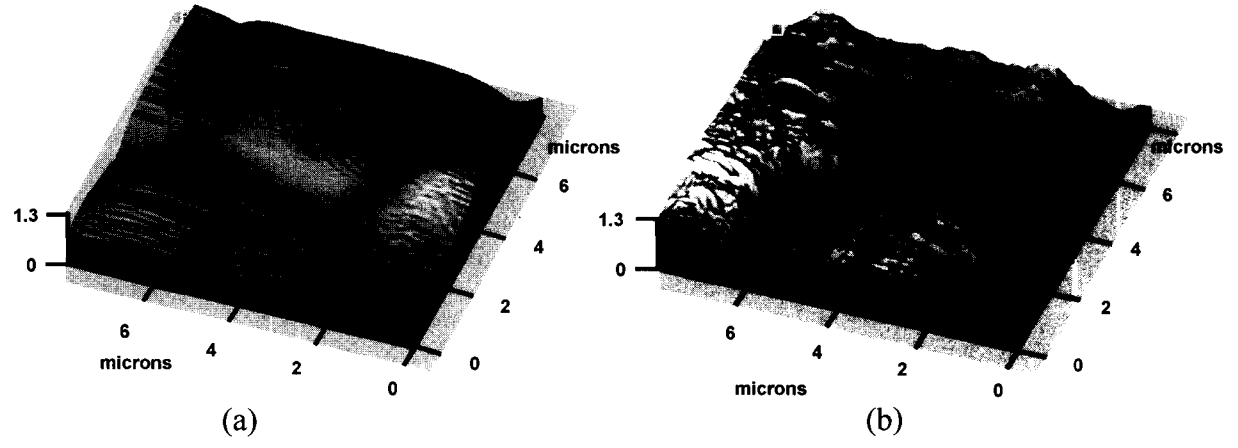
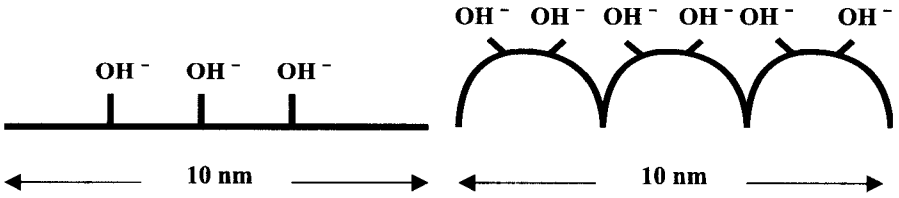


Fig. 9 Representative atomic force images of (a) conventional materials compared to (b) nanophase materials (specifically, titania is depicted here). (From Ref. [35].)



(a) Conventional (grain sizes greater than 100 nm)

(b) Nanophase (grain sizes less than 100 nm)

Fig. 10 Special surface properties of nanophase materials pertinent for protein interactions. Compared to (a) conventional materials, (b) nanophase materials possess a greater number of atoms at the surface, higher surface areas, and less acidic OH^- groups (because of an increase in electron delocalization) in the hydroxide layer. (From Refs. [1] and [7].)

cumulative adsorption of proteins contained in serum was significantly higher on smaller, nanometer grain-sized ceramics.[36] In particular, the interaction of four proteins (fibronectin, vitronectin, laminin, and collagen—proteins all known to enhance osteoblast function)[36] increased greatly on nanophase compared with conventional ceramics.[36] For example, by just decreasing grain size to below 100 nm, select competitive vitronectin adsorption increased by 10% on alumina formulations. These studies demonstrated that initial enhanced calcium adsorption on nanophase ceramics was a key mechanism for increased protein interactions.[36]

Calcium-mediated protein adsorption affected the orientation of proteins adsorbed on nanophase ceramics as well.[40] Specifically, a novel adaptation of the standard surface-enhanced Raman scattering (SERS) technique provided evidence of increased unfolding of the aforementioned proteins adsorbed on nanophase compared with conventionally grain-sized ceramics (Fig. 11).[40] Unfolding of these proteins promoted the availability of specific cell-adhesive epitopes (such as the amino acid sequence Arg–Gly–Asp) that undoubtedly increased bone cell adhesion and function (this interaction is depicted in Fig. 8); evidence supporting this claim was further provided by these investigators through competitive cell adhesion inhibition studies.[40] These studies were the first to demonstrate that by decreasing grain size (or, in effect, decreasing the size of surface features) to below 100 nm (thus becoming closer to the fundamental size of proteins), interactions of proteins were altered in ways previously unobserved on conventional formulations.

It is interesting to consider what properties of nanophase ceramics manipulated protein adsorption and orientation to benefit new bone formation needed for orthopedic implant applications. Select adsorption of proteins important for bone cell function was enhanced by simply decreasing the grain size of the ceramics into the nanometer regime. Equally as important, the tertiary structure of the proteins (namely, the positioning of the Arg–Gly–Asp amino acid sequence in the interior of some of these proteins)[57] was manipulated and partially unfolded to expose cell-adhesive epitopes. Because of these novel observations, the investigators characterized several key material properties of nanophase ceramics to find some insights into these interactions.

One may speculate that increased surface area (up to 30% for alumina)[35] of nanophase materials is the only property leading to increased adsorption of proteins. However, although this property certainly holds promise for biofiltration, bioseparation applications, etc., literature reports have normalized increased protein adsorption to the increased surface area observed on nanophase ceramics, and have still found enhanced protein adsorption.[36] Thus it is clear that the greater surface area that results from nanostructured surface features compared with conventionally structured surface features is not the only property promoting interactions with proteins.

Increased protein adsorption may be directly related to the reported increased surface wettability of nanophase over conventional ceramics.[35] It was previously suspected that nanophase ceramics have much greater surface reactivity because of an increased number of atoms at the surface, greater amounts of grain boundaries at the surface, and higher proportions of edge sites (Fig. 10).[7] This was confirmed by studies that provided evidence of aqueous contact angles three times smaller than when alumina grain size was reduced from 167 to 24 nm.[35] As previously discussed, because proteins assume a tertiary structure with mostly hydrophilic amino acid residues on the exterior of globular shapes, a material surface with increased wettability (or hydrophilic) properties should increase interactions with hydrophilic regions of proteins to, subsequently, enhance unfolding of that protein to expose interior amino acids important for cell adhesion.

In addition, large topographical differences resulting from grain size as well as pore size differences were found between nanophase and conventional ceramics as expected (Fig. 9).[35] Specifically, surface roughness increased by 35–50% on nanophase compared with conventional grain size ceramics.[35] These investigators

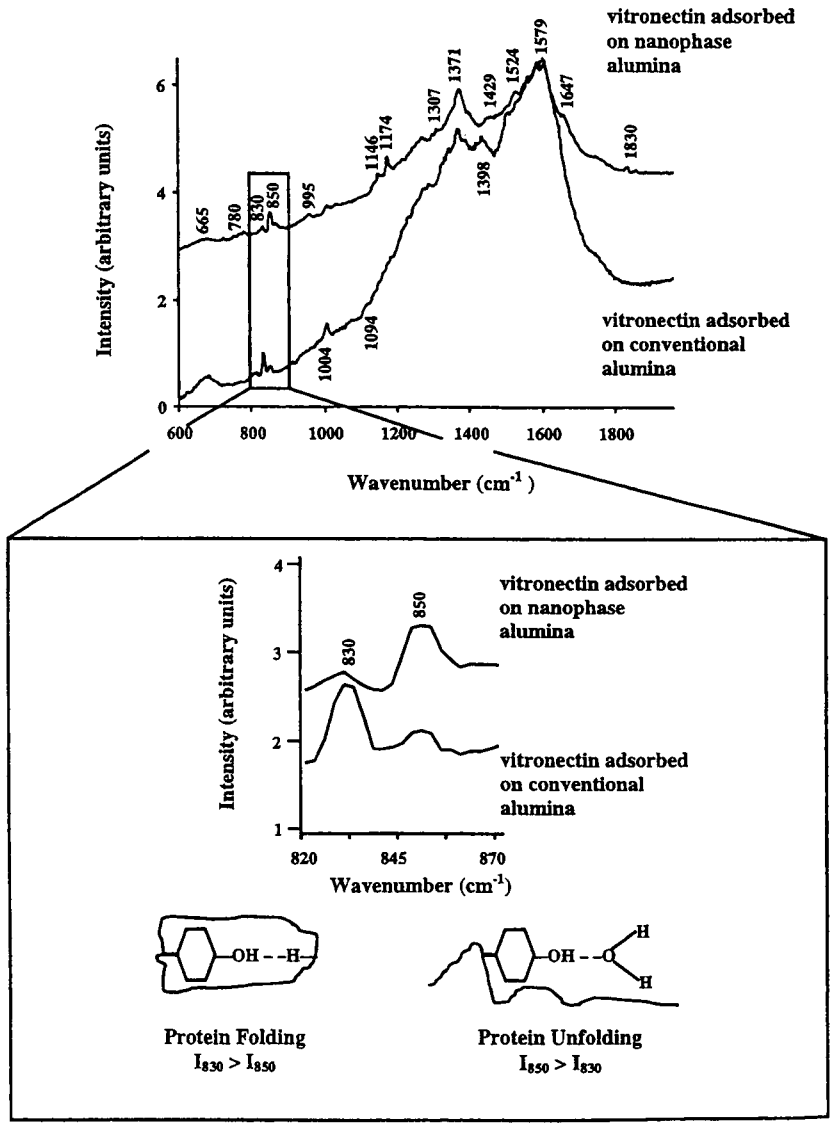


Fig. 11 Surface-enhanced Raman scattering technique provided evidence of increased unfolding of the proteins (specifically, vitronectin in this example) adsorbed on nanophase compared to conventional grain-sized ceramics. (From Ref. [40].)

hypothesized that, for the first time, because of the ability to create surface topographies with features that approximate the size of proteins, extreme control over protein orientation can result. For example, because of protein stereochemical structure and ceramic pore size dimensions (in angstrom regime compared with micron regime for nanophase and conventional ceramics, respectively), many of these nanostructured proteins important for bone cell function adsorbed in preferential orientations to smaller pores sizes in nanophase ceramics.[36,42] Some of the larger proteins contained in blood plasma (such as albumin not important for bone cell function) were preferentially sterically excluded from angstrom-sized pores of nanostructured ceramics.[36,42] In addition, variations in ceramic surface topography on the same order of magnitude as the size of proteins influenced adsorbed protein orientation, and thus the availability of select amino acid sequences, to promote bone-forming cell function. Because of protein dimension in the nanometer regime, through the use of nanostructured topographies, scientists can now modify a surface to control and manipulate adsorbed protein conformation; this is, most likely, the largest unexplored and promising potential at the intersection of protein interactions and nanophase materials. These studies demonstrate that the reported high surface reactivity nanophase materials have

in enhancing catalytic applications can be transferred to benefit applications involving protein interactions.

CONCLUSION

Nanostructured materials provide alternatives not yet fully explored for controlling interactions with proteins. Because technologies now exist to manipulate materials at the atomic, molecular, and supramolecular levels, surfaces can be designed at a dimension similar to that of proteins—the nanometer level. Moreover, in conjunction with this promise, nanophase materials have special surface properties because of an inherent increased number of atoms at the surface, large surface areas, higher proportions of edge sites, and a greater number of material defects (such as grain boundaries for ceramics) at the surface. When these concepts are fully appreciated and realized, it will be possible to design highly engineered surfaces for improving implantable devices, filtration systems, bioseparations, bioMEMS, etc. Such advances were previously unimaginable with conventional materials. Fundamental knowledge of protein dimension, structure, and organization is essential toward this goal. As the disciplines of protein biology and nanophase material science develop and mature, the design criteria mentioned in this chapter for controlling protein interactions will be expanded and refined. Undoubtedly, nanophase materials have the potential to become the next generation-of-choice proactive materials for innovative biotechnology and biomedical applications that could have profound impact in many diverse fields that involve interactions with proteins.

REFERENCES

1. Baraton, M.I.; Chen, X.; Gonsalves, K.E. FTIR study of nanostructured alumina nitride powder surface: Determination of the acidic/basic sites by CO, CO_2, and acetic acid adsorptions. Nanostruct. Mater. **1997**, *8*, 435.
2. Bohn, R.; Haubold, R.; Birringer, R.; Gleiter, H. Nanocrystalline intermetallic compounds—An approach to ductility. Scr. Metall. Mater. **1991**, *25*, 811.
3. Carry, C.; Mocellin, A. Structural superplasticity in single phase crystalline ceramics. Ceram. Int. **1987**, *13* (2), 89–98.
4. Catledge, S.; Vohra, Y. Effect of nitrogen addition on the microstructure and mechanical properties of diamond films grown using high-methane concentrations. J. Appl. Phys. **1999**, *86* (1), 698–700.
5. Ciftcioglu, M.; Mayo, M.J. Processing of Nanocrystalline Ceramics. In *Superplasticity in Metals, Ceramics, and Intermetallics Symposium Proceedings*; Mayo, M.J., Kobayashi, M., Wadsworth, J., Eds.; Materials Research Society: Pittsburgh, 1990; 77–86.
6. Cui, Z.; Hahn, H. Tensile deformation of nanostructured TiO_2 at low temperatures. Nanostruct. Mater. **1992**, *1*, 419.
7. Klabunde, K.J.; Strak, J.; Koper, O.; Mohs, C.; Park, D.; Decker, S.; Jiang, Y.; Lagadic, I.; Zhang, D. Nanocrystals as stoichiometric reagents with unique surface chemistry. J. Phys. Chem. **1996**, *100*, 12141.
8. Mayo, M.; Siegel, R.W.; Liao, Y.X.; Nix, W.D. Nanoindentation of nanocrystalline ZnO. J. Mater. Res. **1992**, *7*, 973.
9. Mayo, M.; Siegel, R.W.; Narayanasamy, A.; Nix, W.D. Mechanical properties of TiO_2 as determined by nanoindentation. J. Mater. Res. **1990**, *5*, 1073.
10. Nieman, G.W. Processing and Mechanical Behavior of Nanocrystalline Cu, Pd, and Ag. Ph.D. Thesis; Northwestern University, 1991.
11. Nieman, G.W.; Weertman, J.R.; Siegel, R.W. Microhardness of nanocrystalline palladium and copper produced by inert gas condensation. Scr. Metall. **1989**, *23*, 2013.
12. Nieman, G.W.; Weertman, J.R.; Siegel, R.W. Mechanical behavior of nanocrystalline metals. J. Mater. Res. **1991**, *6*, 1012.
13. Nieman, G.W.; Weertman, J.R.; Siegel, R.W. Mechanical Behavior of Nanocrystalline Cu and Pd. In *Microcomposites and Nanophase Materials*; Van Aken, D.C., Ed.; TMS: Warrendale, PA, 1991; 15.
14. Siegel, R.W. Nanophase Materials. In *Encyclopaedia of Applied Physics*; VCH Publishers, Inc.: New York, NY, 1994; Vol. 11, 173–199.
15. Siegel, R.W. Creating nanophase materials. Sci. Am. **1996**, *275*, 42.
16. Siegel, R.W.; Fougere, G.E. Mechanical Properties of Nanophase Materials. In *Nanophase Materials: Synthesis–Properties–Applications*; Hadjipanayis, G.C., Siegel, R.W., Eds.; Kluwer: Dordrecht, 1994.
17. Siegel, R.W.; Fougere, G.E. Mechanical properties of nanophase metals. Nanostruct. Mater. **1995**, *6*, 205.
18. Siegel, R.W.; Fougere, G.E. Grain size dependent mechanical properties on nanophase materials. Mater. Res. Soc. Symp. Proc. **1995**, *362*, 219.
19. Toprani, N.; Catledge, S.; Vohra, Y. Interfacial adhesion and toughness of nanostructured diamond coatings. J. Mater. Sci. **2000**, *15* (5), 1052–1055.
20. Weertman, J.R.; Farkas, D.; Hemker, K.; Kung, H.; Mayo, M.; Mitra, R.; van Swygenhoven, H.

Structure and mechanical behavior of bulk nanocrystalline materials. MRS Bull. **1999**, *24* (2), 44–50.
21. Wu, S.J.; DeJong, L.C.; Rahaman, M.N. Sintering of nanophase γ-Al_2O_3 powder. J. Am. Ceram. Soc. **1996**, *79*, 2207.
22. Barringer, E.A.; Bowen, H.K. Formation, packing, and sintering of monodispersed TiO_2 powders. Commun. Am. Ceram. Soc. **1982**, *12*, C199.
23. Patrick, G.L. *An Introduction to Medicinal Chemistry*; Oxford University Press: New York, 1995; 15–26.
24. Suckling, K.E.; Suckling, C.J. *Biological Chemistry*; Cambridge University Press: New York, 1980.
25. Horbett, T.A. Proteins: Structure, Properties and Adsorption to Surfaces. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, A.S., Lemmons, J.E., Eds.; Academic Press: New York, 1996; 133–140.
26. Eisenberg, D. Three-dimensional structure of membrane and surface proteins. Annu. Rev. Biochem. **1984**, *53*, 595–623.
27. Bull, H.B.; Breese, K. Surface tension of amino acid solutions: A hydrophobicity scale of the amino acid residues. Arch. Biochem. Biophys. **1974**, *161*, 665–670.
28. Peters, T. Serum Albumin. In *Advances in Protein Chemistry*; Anfinsen, C.B., Edsall, J.T., Richards, F.M., Eds.; Academic Press: New York, 1985; Vol. 37, 161–245.
29. Stryer, L. *Biochemistry,* 2nd Ed.; W. H. Freeman: San Francisco, 1981.
30. Loeb, W.F.; Mackey, W.F. A "cuvette method" for the determination of plasma fibrinogen. Bull. Am. Soc. Vet. Clin. Pathol. **1972**, *1*, 5–8.
31. Norde, W.; Lyklema, J. Why proteins prefer interfaces. J. Biomater. Sci., Polym. Ed. **1991**, *2*, 183–202.
32. Yutani, K.; Ogasahara, K.; Tsujita, T.; Sugino, Y. Dependence of conformational stability on hydrophobicity of the amino acid residue in a series of variant proteins substituted at a unique position of tryptophan synthase alpha subunit. Proc. Natl. Acad. Sci. U. S. A. **1987**, *84*, 4441–4444.
33. Webster, T.J.; Siegel, R.W.; Bizios, R. An In Vitro Evaluation of Nanophase Alumina for Orthopaedic/Dental Applications. In *Bioceramics 11: 11th International Symposium on Ceramics in Medicine*; LeGeros, R.Z., LeGeros, J.P., Eds.; 1998; 273.
34. Webster, T.J.; Siegel, R.W.; Bizios, R. Design and evaluation of nanophase alumina for orthopaedic/dental applications. Nanostruct. Mater. **1999**, *12*, 983.
35. Webster, T.J.; Siegel, R.W.; Bizios, R. Osteoblast adhesion on nanophase ceramics. Biomaterials **1999**, *20*, 1221.
36. Webster, T.J.; Ergun, C.; Doremus, R.H.; Siegel, R.W.; Bizios, R. Specific proteins mediate enhanced osteoblast adhesion on nanophase ceramics. J. Biomed. Mater. Res. **2000**, *51* (3), 475.
37. Webster, T.J.; Siegel, R.W.; Bizios, R. Enhanced functions of osteoblasts on nanophase ceramics. Biomaterials **2000**, *21*, 1803.
38. Webster, T.J.; Siegel, R.W.; Bizios, R. Enhanced Surface and Mechanical Properties of Nanophase Ceramics for Increased Orthopaedic/Dental Implant Efficacy. In *Bioceramics 13: 13th International Symposium on Ceramics in Medicine Conference Proceedings*; Giannini, S., Moroni, A., Eds.; 2000; 321.
39. Webster, T.J. Nanophase Ceramics: The Future Orthopedic and Dental Implant Material. In *Advances in Chemical Engineering*; Ying, J.Y., Ed.; Academic Press: New York, 2001; Vol. 27, 125–166.
40. Webster, T.J.; Schadler, L.S.; Siegel, R.W.; Bizios, R. Mechanisms of enhanced osteoblast adhesion on nanophase alumina involve vitronectin. Tissue Eng. **2001**, *7* (3), 291–301.
41. Webster, T.J.; Ergun, C.; Doremus, R.H.; Siegel, R.W.; Bizios, R. Enhanced functions of osteoclast-like cells on nanophase ceramics. Biomaterials **2001**, *22* (11), 1327–1333.
42. Webster, T.J.; Siegel, R.W.; Bizios, R. Nanoceramic surface roughness enhances osteoblast and osteoclast functions for improved orthopaedic/dental implant efficacy. Scr. Mater. **2001**, *44*, 1639–1642.
43. Horbett, T.A. Principles underlying the role of adsorbed plasma proteins in blood interactions with foreign materials. Cardiovasc. Pathol. **1993**, *2*, 137S–148S.
44. Horbett, T.A. Techniques for Protein Adsorption Studies. In *Techniques of Biocompatibility Testing*; Williams, D.F., Ed.; CRC Press: Boca Raton, FL, 1986; 183–214.
45. Horbett, T.A.; Brash, J.L. Proteins at Interfaces: Current Issues and Future Prospects. In *Proteins at Interfaces: Physiochemical and Biochemical Studies*; Horbett, T.A., Brash, J.L., Eds.; American Chemical Society: Washington, DC, 1987; Vol. 343, 1–33.
46. Horbett, T.A. Protein Adsorption on Biomaterials. In *Biomaterials: Interfacial Phenomena and Applications*; Cooper, S.L., Peppas, N.A., Eds.; American Chemical Society: Washington, DC, 1982; Vol. 199, 233–244.
47. Andrade, J.D. Principles of Protein Adsorption. In *Surface and Interfacial Aspects of Biomedical*

Polymers; Andrade, J., Ed.; Plenum Publ.: New York, 1985; 1–80.

48. Schakenraad, J.M. Cell: Their Surfaces and Interactions with Materials. In *Biomaterials Science: An Introduction to Materials in Medicine*; Ratner, B.D., Hoffman, A.S., Schoen, A.S., Lemmons, J.E., Eds.; Academic Press: New York, 1996; 141–147.
49. Sinha, R.K.; Tuan, R.S. Regulation of human osteoblast integrin expression by orthopedic implant materials. Bone **1996**, *18*, 451–457.
50. Brunette, P.M. The Effect of Surface Topography of Cell Migration and Adhesion. In *Surface Characterization of Biomaterials: Progress in Biomedical Engineering*; Ratner, B.D., Ed.; Elsevier: New York, 1988; Vol. 6, 203–217.
51. Davies, J.E. The Importance and Measurement of Surface Charge Species in Cell Behavior at the Biomaterial Interface. In *Surface Characterization of Biomaterials: Progress in Biomedical Engineering*; Ratner, B.D., Ed.; Elsevier: New York, 1988; Vol. 6, 219–234.
52. Luck, M.; Paulke, B.R.; Schroder, W.; Blunk, T.; Muller, R.H. Analysis of plasma protein adsorption on polymeric nanoparticles with different surface characteristics. J. Biomed. Mater. Res. **1998**, *39*, 478–485.
53. Curtis, A.; Wilkinson, C. Review: Topographical control of cells. Biomaterials **1997**, *18* (24), 1573–1583.
54. Degasne, I.; Basle, M.F.; Demais, V.; Hure, G.; Lesourd, M.; Grolleau, B.; Mercier, L.; Chappards, D. Effects of roughness fibronectin and vitronectin on attachment, spreading, and proliferation of human osteoblast-like cells (Saos-2) on titanium surfaces. Calcif. Tissue Int. **1999**, *64* (6), 499–507.
55. Dalton, B.A.; McFarland, C.D.; Gengenbach, T.R.; Griesser, H.J.; Steele, J.G. Polymer surface chemistry and bone cell migration. J. Biomater. Sci., Polym. Ed. **1995**, *9* (8), 781–799.
56. Lopes, M.A.; Monteiro, F.J.; Santos, J.D.; Serro, A.P.; Saramago, B. Hydrophobicity, surface tension, and zeta potential measurements of glass-reinforced hydroxyapatite composites. J. Biomed. Mater. Res. **1999**, *45* (4), 370–375.
57. Ayad, S.; Boot-Handford, R.; Humphries, M.J.; Kadler, K.E.; Shuttleworth, A. *The Extracellular Matrix Factsbook*; Academic Press: San Diego, CA, 1994.

Quantum Dot Arrays: Electromagnetic Properties

Sergey A. Maksimenko
Gregory Ya. Slepyan
Belarus State University, Minsk, Belarus

INTRODUCTION

A fundamental breakthrough in semiconductor device physics is connected with the recent progress in the synthesis of sheets of nanoscale narrow-gap insertions in a host semiconductor—quantum dots (QDs). The key peculiarity of QDs is related to the spatial confinement of charge carriers and their discrete energy levels determined by QD size and shape. The shape is dictated by the growth kinetics and the parameters of the materials governing the thermodynamics of growth. The large body of recent results on the physical properties of QDs and their utilization for the QD laser design has been discussed in Ref. [1].

Apart from charge carrier confinement, there exists a class of effects governed by the intrinsic spatial inhomogeneity of the QD heterostructures. Transparent and dissipative heterogeneous media with small inclusions of one material into another material, conventionally referred to as composite materials, exhibit, in general, new mechanical, electronic, as well as optical properties, which are not inherent to each individual component.[2] In the case of QD heterostructures, a conventional picture is modified because of specific properties of excitons, coupled electron–hole states, which define the QD response. First, excitonic composite is constituted by resonant particles and, consequently, is characterized by resonant response; moreover, inverse population is possible owing to the discrete energy spectrum of excitons. Thus a QD-based composite is a resonant active system. Another specific property of QD composites appears owing to the quantum nature of excitons: the exciton Bohr radius a_B can either exceed the QD linear extension (strong confinement regime), or be much less (weak confinement regime). In the latter case, often realized in experiments, the QD electromagnetic response becomes nonlocal[3] (i.e., constitutive relations for polarization of the QD medium take the form of integral operators). Electromagnetic properties of QD-based composites are the focus of the present paper. Consideration is based on earlier published articles of various authors.[4–10]

BASIC RELATIONS

Effective Permittivity Tensor for Three-Dimensional Structures

Electromagnetic properties of composites are usually modelled in the framework of the effective medium approach,[2] which implies electromagnetic field averaging over material inhomogeneities. Thus a homogeneous medium with effective constitutive parameters (ECPs), such as conductivity, susceptibility, and permittivity, instead of a composite, is being considered. The effective parameters are expressed in terms of the generic and the geometrical parameters of the inclusions and the host medium. The general approach for estimating the effective parameters of a composite material is as follows: First, the field scattered by a single inclusion in the host medium is found; then, the scattering contributions from all inclusions are summed and averaged over a vanishingly small region. For this approach to hold, all inclusions must be electrically small (i.e., their linear size must be small compared with the wavelength). Quantum dot-based structures completely satisfy that condition in the visible range.

There are a large number of different modifications of the effective medium theory. Among them, we choose the Maxwell Garnett approach as it is based on rigorous solutions of the integral equations of macroscopic electrodynamics[11] for composites with small volume fractions of inclusions (f_V<0.4–0.5) (i.e., a weak modification of the electronic spectrum and the gain of QD ensemble are assumed as compared with that of individual QDs). In the framework of the Maxwell Garnett approach, rigorous derivation of the effective permittivity tensor has been first presented in Refs. [12] and [13] (also Ref. [2]) where it has been shown that a composite medium comprising a regular ensemble of uniform-sized, electrically small, dielectric inclusions dispersed in a host dielectric material is characterized by the effective permittivity tensor as follows:

$$\hat{\epsilon}(\omega) = \varepsilon_h + f_V \varepsilon_h \hat{\alpha}(\omega)[1 + f_V \hat{\delta}\hat{\alpha}(\omega)]^{-1} \quad (1)$$

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013524

where $\hat{\alpha}(\omega)$ is the polarizability tensor of a single QD and $\hat{\delta}$ is the lattice tensor completely determined by the array geometry. Furthermore, the dielectric constant of the host material ε_h is assumed to be real, frequency-independent, and equal to the permittivity of the QD material far away from the exciton resonance.

Eq. 1 states that the QD ensemble comprises an optically anisotropic medium even if both the QD and host materials are isotropic. Thus we predict electromagnetic anisotropy of the gain in QD arrays because of the diffraction of the electromagnetic field by inclusions. The diffraction manifests itself in two ways. First, the effective permittivity tensor depends on the array geometry involving the electromagnetic interaction between inclusions. Second, the polarizability of a single QD depends on its shape because of the depolarizing field. Assuming the linear size of the lattice elementary cell to be much less than the wavelength, it has been shown that components of the lattice tensor are given by integrals as follows:

$$\delta_{ij} = \frac{1}{4\pi}\int_{\Omega}\frac{r_i r_j}{r^5}\mathrm{d}^3 r \text{ if } i \neq j \quad \text{and}$$
$$\delta_{ii} = \frac{1}{4\pi}\int_{\Omega}\frac{3r_i^2 - r^2}{r^5}\mathrm{d}^3 r \tag{2}$$

where Ω is the elementary cell volume, $\boldsymbol{r}$ is the radius vector of a point inside the elementary cell, $r=|\boldsymbol{r}|$; indices stand for Cartesian components. Note that lattice tensor 2 is found without any reference to specific properties of QDs as quantum-mechanical objects.

Effective Boundary Conditions for Planar Structures

In many cases, a planar array of QDs with intrinsic two-dimensional (2-D) periodicity of characteristic period much less than the optical wavelength can be treated as a more adequate and realistic model.[1] Below we present a general method for the evaluation of the electromagnetic response of planar arrays of QDs. This method, conventionally referred to as the effective boundary condition (EBC) method, has been originally developed for microwaves and antenna theory,[14–16] and has found a wide application in these fields. Similar approaches have also been developed in acoustics, hydrodynamics, and elasticity theory. Recently, the EBC method has been extended to low-dimensional nanostructures, such as quantum wells,[17,18] carbon nanotubes,[19,20] and semicontinuous metal films.[21] The basic idea of the EBC method is that a smooth homogeneous surface is considered instead of the initial structure, and appropriate EBCs for the electromagnetic field are stated for this surface. These conditions are chosen in such a way that the spatial structures of the electromagnetic field, because of an effective current induced on the homogeneous surface, and the electromagnetic field of the real current in the initial structure turn out to be identical at some distance away from the surface. Material characteristics of the structure as well as its geometrical parameters are included in coefficients of EBCs. In essence, the EBC method is a modification of the effective medium theory as applied to 2-D confined structures. A detail description of the EBC method as applied to planar QD structures has been presented in Ref. [7].

Effective boundary conditions, in the form of two-sided impedance boundary conditions, were first formulated for the linear electrodynamics of single-shell and multishell carbon nanotubes.[19] Using this analysis, the EBC method has been extended to planar QD structures and corresponding EBCs have been derived:[7]

$$\mathbf{n}\times\mathbf{n}\times(\mathbf{H}^{\mathrm{I}} - \mathbf{H}^{\mathrm{II}}) = -\frac{2\pi}{c}\mathbf{n}\times\hat{\boldsymbol{\sigma}}(\mathbf{E}^{\mathrm{I}} + \mathbf{E}^{\mathrm{II}})$$
$$\mathbf{n}\times(\mathbf{E}^{\mathrm{I}} - \mathbf{E}^{\mathrm{II}}) = -\xi\mathbf{n}\times\nabla[\mathbf{n}\cdot(\mathbf{E}^{\mathrm{I}} + \mathbf{E}^{\mathrm{II}})] \tag{3}$$

where $\mathbf{n}\|\mathbf{e}_z$ is the normal to the QD layer, and $\hat{\sigma}$ and ξ are coefficients completely determined by the shape and size of QDs. Nonzero components of the tensor $\hat{\sigma}$, σ_{ij} $(i,j=x,y)$, constitute 2×2 surface conductivity tensor $\hat{\sigma}_{\|}$. For quadratic lattice with period d, it can easily be found that:

$$\hat{\boldsymbol{\sigma}}_{\|} = i\frac{\varepsilon_h\omega}{\mathrm{d}^2}\left[\hat{\mathbf{I}}_{\|} + \frac{\delta_x}{\mathrm{d}^2}\hat{\alpha}_{\|}\right]^{-1}, \quad \xi = \frac{2\pi\alpha_{zz}}{\mathrm{d}^2 + \delta_z\alpha_{zz}} \tag{4}$$

Here $\hat{\mathbf{I}}_{\|}$ is the 2×2 unit tensor, and $\hat{\boldsymbol{\alpha}}_{\|}$ is given by the in-plane components α_{ij} $(i,j=x,y)$ of the QD polarizability tensor. Coefficients $\delta_{x,z}$ are derived from Eq. 2 for planar quadratic elementary cell.[7]

Eq. 3 constitutes the complete system of EBCs for electromagnetic field in low-dimensional nanostructures. They have been obtained in an ordinary way, by the averaging of a microscopic field over a physically infinitesimal volume. The technique of macroscopic averaging is similar to that which introduces the constitutive parameters for bulk media, but differs in that the averaging occurs in boundary conditions rather than in field equations. Correspondingly, the averaging was carried out over the 2-D surface but not over the three-dimensional (3-D) spatial element. Thus in electrodynamics of low-dimensional structures, EBCs play the same role as constitutive relations in the electrodynamics of bulk media. A theory is applicable in both strong and weak confinement regimes. Although EBCs have been derived for 2-D periodical structures with quadratic lattices, they keep validity for the arbitrary configuration of elementary cell and for planar layers with random

distributions of QDs. The derived EBCs turn out to be analogous to well-known boundary conditions for quantum wells (QWs)[17] if spatial dispersion in the latter can be neglected (i.e., in the limit of infinitely large exciton mass). Thus a planar layer comprising a 2-D array of QDs can be treated as an effective QW.

POLARIZABILITY OF A SINGLE QUANTUM DOT

For practical utilization of the derived ECPs (Eq. 1) or EBCs (Eq. 3), the polarizability tensor of an isolated QD requires to be known. For the simplest configuration of QDs (spheres, disks), this tensor can be found analytically whereas direct numerical simulation is required for more complicated configurations. A conventional phenomenological model of gain in a QD is based on a semiclassical theory of two-level systems, which gives the equation of motion for the mean polarization $\mathbf{P}$ caused by transitions between levels:[22]

$$\left[\frac{\partial^2}{\partial t^2} + \frac{2}{\tau}\frac{\partial}{\partial t} + \omega_0^2\right]\mathbf{P} = -\frac{\omega_0}{2\pi\varepsilon_h}\hat{\mathbf{g}}_0\mathbf{E}_L \tag{5}$$

Here ω_0 is the resonant frequency of the transition, τ is the exciton dephasing time, and $\mathbf{E}_L$ is the macroscopic local field inside the QD. The tensor $\hat{\mathbf{g}}_0$ is a phenomenological parameter expressed in terms of the QD dipole moment $\boldsymbol{\mu}$ diadic:

$$\hat{\mathbf{g}}_0 = \mp\frac{4\pi}{\hbar V}\boldsymbol{\mu}\boldsymbol{\mu}$$

Here, V is the QD volume, and upper and bottom signs correspond to ground and excited states, respectively, so that $\|\hat{\mathbf{g}}_0\|>0$ in an inverted medium. If QD comprises an isotropic material, orientational averaging leads to the change $\boldsymbol{\mu}\boldsymbol{\mu}\rightarrow|\boldsymbol{\mu}_0|^2\hat{\mathbf{I}}/3$, where $\boldsymbol{\mu}_0$ is the matrix element of the dipole moment of the corresponding bulk sample.

For time-harmonic fields, the solution of Eq. 5 in the vicinity of the resonance gives the well-known Lorentz contribution to medium polarizability:

$$\hat{\boldsymbol{\alpha}}(\omega) = \frac{1}{\varepsilon_h}\frac{\hat{\mathbf{g}}_0}{\omega-\omega_0+i/\tau} \tag{6}$$

This dependence on the isotropic $\hat{\mathbf{g}}_0$ is commonly used as a phenomenological model of the dispersion and the gain of a single QD. However, relation (6) assumes an infinite, isotropic, homogeneous medium and can serve only as a rough approximation to the case under consideration: the polarizability of a single QD turns out to be distinct from that given by Eq. 6. Indeed, the local field $\mathbf{E}_L$ inside the QD is different from the external acting field $\mathbf{E}_0$ (Fig. 1). The depolarization field determines this difference by:[23]

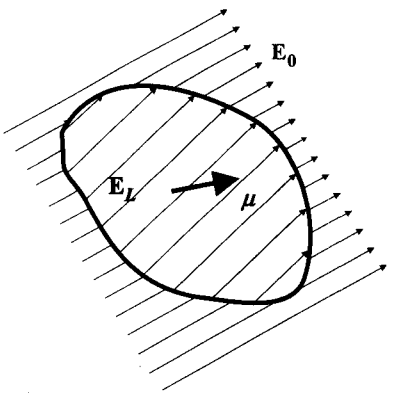


Fig. 1 Schematic picture of correlation between acting and local fields in electrically small scatterer. The local field is assumed to be homogeneous.

$$\mathbf{E}_L = \mathbf{E}_0 - 4\pi\hat{\mathbf{N}}\mathbf{P} \tag{7}$$

where $\hat{\mathbf{N}}$ is the depolarization tensor. This tensor is symmetrical[24] (also Ref. [10]) and depends only on the shape of the scattering object:

$$N_{\alpha\beta} = -\frac{1}{4\pi V}\int_V\int_V\frac{\partial^2}{\partial r_\alpha\partial r_\beta}\frac{1}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}d^3\mathbf{r}'$$

Here $r_{\alpha,\beta}=x,y,z$. For a sphere, $\hat{\mathbf{N}}=\hat{\mathbf{I}}/3$; for an ellipsoid, this tensor is diagonal on a basis related to the ellipsoid's axes.[23] For more complex objects, its elements can be found only numerically.

Substitution of Eq. 7 into Eq. 5 results in a set of equations for three coupled oscillators, which describe the tensorial polarizability of QD in the vicinity of resonance:

$$\hat{\alpha}(\omega) = \frac{1}{\varepsilon_h}\frac{1}{(\omega-\omega_0+i/\tau)\hat{\mathbf{I}}+\hat{\mathbf{g}}_0\hat{\mathbf{N}}/\varepsilon_h}\hat{\mathbf{g}}_0 \tag{8}$$

In this equation, we neglect the linewidth change provided by depolarization.[15] Thus one can conclude that the QD's shape reflects itself as a fine structure of the resonance, which itself is a superposition of three bands with frequencies $\omega_0-\Delta\omega_j$, where $\Delta\omega_j$ are the eigenvalues of the inner tensor product $\hat{\mathbf{g}}_0\hat{\mathbf{N}}/\varepsilon_h$. Note that this diffraction-induced shift is analogous to that which occurs in ferromagnetic resonance, where the resonance frequencies of small-sized grains and infinite continuous media are different because of the effects of demagnetization.[23] Depolarization shift has been ignored in Ref. [25] under the construction of the electromagnetic response theory of QD arrays. For spherical isotropic QDs, the tensors $\hat{\mathbf{N}}$ and

$\hat{\mathbf{g}}_0$ are isotropic and the fine structure manifests itself as a polarization-independent shift of the resonance line.[26,27] Let us estimate the shift using well-known data for QD characteristics and relations for the exciton radiative lifetime in a medium $\hbar/\tau_{rad}=4k_1^3|\boldsymbol{\mu}_0|^2/3$, where $k_1=\sqrt{\varepsilon_h}\omega_0/c$. Then, the depolarization shift can be written by:

$$\Delta\omega = \pm\frac{4\pi}{9\hbar V}|\boldsymbol{\mu}_0|^2 = \pm\frac{1}{\tau_{rad}}\frac{\pi}{Vk_1^3} \tag{9}$$

For GaAs spherical QDs with radius in the range R=2–4 nm, dielectric constant ε_h=12.25, and radiative lifetime $\tau_{rad}\cong$1 nsec,[1] at the wavelength λ=1.3 μm, formula 9 gives $\hbar\Delta\omega\sim$10–1 meV, respectively. This estimate correlates well with the theoretical predictions given in Refs. [4,5,28]. Note that the Bohr radius for such QDs is about 10 nm,[1] so that the strong confinement approximation used above is valid. For experimental detection of the predicted fine structure, the value $\Delta\omega$ must exceed the linewidth provided by radiative decay and homogeneous broadening because of dephasing. From Eq. 9, it follows that the inequality $\Delta\omega\gg 1/\tau_{rad}$ holds true for any realistic arbitrary-shaped QDs. Analysis shows that the dominant contribution to the homogeneous broadening gives exciton–phonon interactions. Recent low-temperature (T=20–40 K) measurements[29–31] give the homogeneous linewidth ~1–20 μeV. Analogous estimates follow from calculations presented in Ref. [32] at T=77 K. Thus at low temperatures, the predicted value of the shift turns out to be sufficiently large to be measured. At room temperature, the homogeneous linewidth grows up to ~0.2–1 meV.[1,31,32] This magnitude is comparable to the depolarization shift. However, even in that case, local field effects are of importance for adequate prediction of spectral line shape. Note that the optical absorption and gain of an isolated QD could be distinguished owing to the depolarization shift—blue in the former case and red in the latter one.[27]

Physical Interpretation

Physical interpretation of the depolarization field effect can be given by an analogy with the $\mathbf{k}\cdot\mathbf{p}$ theory of bulk crystals[33] utilizing the concept of electron–hole effective mass. For spherical QD, using the standard expression for ω_0 and the given expression for $\Delta\omega$, we obtain:

$$\hbar(\omega_0+\Delta\omega) = \varepsilon_g + \frac{\hbar^2\kappa_{nl}^2}{2R^2M} \pm \frac{|\mu|^2}{R^3} \tag{10}$$

where M is the mass of electron–hole pair in QD, ε_g is the width of the forbidden bandgap, κ_{nl} is the nth root of the Bessel function $J_{l+1/2}(\cdot)$, and indices n and l define the working mode in the oscillator spectrum. The third term in the right-hand part of this equation describes the depolarization field contribution. The right-hand part of the equation can be rewritten as:

$$\varepsilon_g + \hbar^2\kappa_{nl}^2/2R^2M_{eff}$$

with M_{eff} given by:

$$M_{eff} = M\left[1 \pm \frac{|\mu|^2M}{\hbar^2\kappa_{nl}^2R}\right]^{-1}$$

The quantity M_{eff} can be interpreted as effective mass of the electron–hole pair in the QD. Thus electromagnetic effects at the QD boundary (QD depolarization) change the exciton effective mass. Analogous consideration for the case of asymmetrically shaped QDs leads to the tensorial effective mass, which gives rise to the polarization-dependent splitting of the gain band.[4,5]

The Role of Nonlocality

In the weak confinement regime, when the exciton Bohr radius is much less than the QD linear extension, the QD electromagnetic response becomes nonlocal: The constitutive relation for medium polarization takes the form of the integral operator as follows:[34]

$$\mathbf{P}(\mathbf{r}) = \frac{DF(\mathbf{r})}{\omega-\omega_0+i/\tau}\boldsymbol{\Lambda} \tag{11}$$

where

$$\boldsymbol{\Lambda} = \int_V F(\mathbf{r}')\mathbf{E}_L(\mathbf{r}')d^3\mathbf{r}', \quad D = -2|\boldsymbol{\mu}|^2\Phi^2(0)/\hbar \tag{12}$$

The function $F(\mathbf{r})$ is the envelope function of the exciton ground state whereas the function $\Phi(r)=\exp(-r/a_B)/\sqrt{\pi a_B^3}$[35] describes the electron–hole relative motion. Inside, QD retardation can be neglected and thus local and acting fields are coupled by:

$$\mathbf{E}_L(\mathbf{r}) = \mathbf{E}_0(\mathbf{r}) + \nabla\nabla\cdot\int_V\frac{\mathbf{P}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}' \tag{13}$$

Assuming polarization to be constant over the QD volume, Eq. 13 is reduced to relation 7. Eq. 13 allows us to find vector $\boldsymbol{\Lambda}$ by omitting the procedure of evaluation of the local electromagnetic field $\mathbf{E}_L(\mathbf{r})$; to do this, let us multiply this equation by the function $F(\mathbf{r})$ and integrate it over the QD volume. As a result, we obtain:

$$\boldsymbol{\Lambda} \approx F_0\mathbf{E}_0 + \frac{\tilde{g}_0}{\omega-\omega_0+i/\tau}\hat{\mathbf{N}}'\boldsymbol{\Lambda}$$

where

$$F_0 = \int_V F(\mathbf{r}')d^3\mathbf{r}', \quad \tilde{g}_0 = \frac{4\pi\varepsilon_h}{V}F_0^2D \tag{14}$$

and the 3-D tensor $\hat{\mathbf{N}}'$ is given by its components:

$$N'_{\alpha\beta} = -\frac{V}{4\pi F_0^2}\int_V\int_V F(\mathbf{r})F(\mathbf{r}')\frac{\partial^2}{\partial r_\alpha \partial r_\beta} \times \frac{1}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}' \quad (15)$$

In the far zone, the electromagnetic field scattered by QD is characterized by the Hertz potential:

$$\boldsymbol{\Pi}^e = \frac{e^{ik_1 r}}{r}\int_V \mathbf{P}(\mathbf{r}')\mathrm{d}^3\mathbf{r}' = \frac{e^{ik_1 r}}{r}DF_0\boldsymbol{\Lambda}$$

Substituting the above relation for $\boldsymbol{\Lambda}$ into this formula and taking into account the relation:[23]

$$\boldsymbol{\Pi}^e = \frac{V}{4\pi r}e^{ik_1 r}\hat{\boldsymbol{\alpha}}\mathbf{E}_0$$

we come to Eq. 8 for the polarizability tensor of an isolated QD with the change $\hat{\mathbf{g}}_0 \to \tilde{g}_0\hat{\mathbf{I}}$ and $\hat{\mathbf{N}} \to \hat{\mathbf{N}}'$. Thus the special law of nonlocality (Eq. 11) inherent to an isolated QD admits the description of the electromagnetic field scattering by the QD using the polarizability tensor independent of the incident field structure. In other words, the nonlocality changes the values of the polarizability tensor components but does not change the general representation of the scattering operators compared with the strong confinement regime. The above result admits extension of the Maxwell Garnett approach to 3-D composites constituted by QDs in weak confinement regimes.

Let us estimate depolarization shift in the weak confinement regime for a spherical QD of the radius R. Exciton wavefunction for an isolated spherical QD is as follows (e.g., Ref. [26]):

$$F(r) \equiv F_{nlm} = C_{nl}Y_{lm}(\vartheta,\varphi)J_{l+1/2}(\kappa_{nl}\rho/R)/\sqrt{\rho}$$

where $Y_{lm}(\vartheta,\varphi)$ is the spherical harmonics; $J_{l+1/2}(x)$ is the Bessel function; κ_{nl} is its nth root; ρ, ϑ, and φ are the spherical coordinates; and indices n and l define the working mode in the oscillator spectrum. Coefficient $C_{nl} = \sqrt{2}[RJ_{l+3/2}(\kappa_{nl})]^{-1}$ provides orthonormalization of functions F_{nlm}. Furthermore, we restrict ourselves to the case $l=m=0$. By integrating Eqs. 14 and 15, we obtain $F_0{}^2 = 32\pi R^3/\kappa_{n0}^2$, $N_{xx}=N_{yy}=N_{zz}=\kappa_{n0}^2/72$. This allows us to evaluate the depolarization shift of the exciton resonance in the weak confinement regime:

$$\Delta\omega = \pm\frac{4}{3\tau_{\mathrm{rad}}}\frac{\pi}{Vk_1^3}\left(\frac{R}{a_B}\right)^3 \quad (16)$$

This shift is different from that given by Eq. 9 for spherical QDs in the strong confinement regime. Because $V\sim R^3$, formula 16 states the independence of the depolarization shift on the QD radius in the weak confinement regime. Choosing excitonic modes with another set of numbers n,l,m, we obviously obtain another magnitude of components $\hat{\mathbf{N}}'$ and, consequently, another magnitude of the polarizability tensor $\hat{\boldsymbol{\alpha}}$ (Eq. 8). Thus we come to a peculiar property of the weak confinement regime: Depolarization depends on the excitonic mode number. This property follows from the nonlocality of the exciton polarization. Unlike that, in the strong confinement regime, the depolarization tensor $\hat{\mathbf{N}}$ is completely determined by the QD geometry.

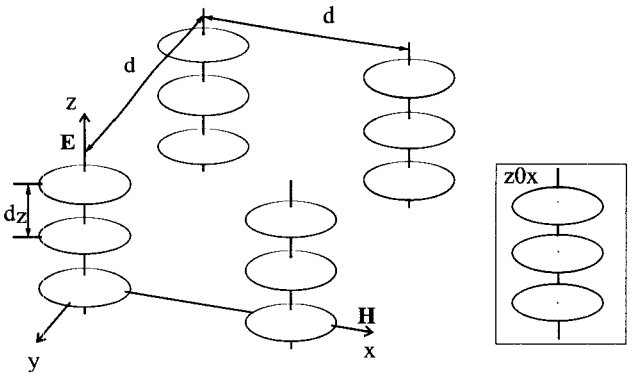


Fig. 2 Array of vertically stacked disklike QDs arranged on a tetragonal lattice.

BIREFRINGENCE IN QUANTUM DOT ARRAYS

In this section, the general consideration developed above applies to the tetragonal lattice composed of certainly shaped identical QDs. The tetragonal lattice serves as a geometrical model for an ensemble of separated QDs vertically stacked with finite period.[1] Let QD be a spheroid with the axis of symmetry directed along the lattice vector $\mathbf{e}_z$ (Fig. 2). This model admits analytical treatment in the framework of the effective medium theory presented above. Indeed, for the chosen QD shape and the lattice geometry, the depolarization tensor $\hat{\mathbf{N}}$ and the lattice tensor $\hat{\boldsymbol{\delta}}$ turn out to be diagonal in the Cartesian basis depicted in Fig. 2,[13] and eigenvalues of these tensors corresponding to the basis vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ are degenerated. Furthermore, the tensor $\hat{\mathbf{g}}_0$ must also show the same properties in these coordinates by reasoning of symmetry. In such a situation, the effective permittivity tensor $\hat{\epsilon}(\omega)$ (Eq. 1) of a composite can be expressed in terms of Cartesian basis diadics by:

$$\hat{\epsilon}(\omega) = \varepsilon_H(\mathbf{e}_x\mathbf{e}_x + \mathbf{e}_y\mathbf{e}_y) + \varepsilon_E\mathbf{e}_z\mathbf{e}_z \quad (17)$$

where

$$\varepsilon_\sigma(\omega) = \varepsilon_h + \frac{f_V \alpha_\sigma(\omega)}{1 + f_V \delta_\sigma \alpha_\sigma(\omega)} \quad (18)$$

and $\sigma = E,H$ refers to light polarized along (E-polarization) or normal (H-polarization) to the z-axis. The polarizability components α_σ follow from Eq. 8 and have the form:

$$\alpha_\sigma(\omega) = \frac{1/\varepsilon_h}{\omega - (\omega_0 - g_0^\sigma N_\sigma/\varepsilon_h) + i/\tau} \quad (19)$$

In accordance with Ref. [23], for spheroid, the depolarization factors N_σ are as follows:

$$N_E = \frac{e^2+1}{e^3}(e - \arctan e), \quad N_H = \frac{1}{2}(1 - N_E) \quad (20)$$

where $e = [a^2/b^2 - 1]^{1/2}$ is the spheroid eccentricity, and a and b are the spheroid semiaxes in the xy-plane and the z-direction, respectively. These formulae hold true for both disklike ($a>b$) and cigarlike ($a<b$) spheroids. Infinite stretching of the spheroids ($a/b \to 0$) results in $N_E \to 0$ and $N_H \to 1/2$, and Eq. 19 reproduces the polarizabilities of cylinders (e.g., Ref. [4]). The geometrical coefficients δ_σ for a tetragonal lattice have been derived in Ref. [13] (also Refs. [4] and [5]):

$$\delta_E = \frac{1}{4\pi}\left[\left(18 - \frac{10}{\beta^2}\right)\arctan\frac{Q}{\beta^2} + 30\frac{1-\beta^2}{\beta^4} \times \left(2\arctan Q - \ln\frac{1+Q}{1-Q}\right)\right]$$

$$\delta_H = \frac{1}{4\pi}\left[\left(18 + \frac{20}{\beta^2} - \frac{30}{\beta^4}\right)\arctan Q + 15\frac{1-\beta^2}{\beta^4}\ln\frac{1+Q}{1-Q}\right] \quad (21)$$

where $\beta = d_z/d$ is the ratio of lattice periods in the z-direction and in the x-direction or y-direction, $Q = \beta/\sqrt{2+\beta^2}$. The above expressions are not valid for $\beta \ll 1$ or $\beta \gg 1$. For a cubic lattice, $\beta = 1$ and the electromagnetic interaction of QDs becomes isotropic: $\delta_E = \delta_H = 1/3$. Thus for the chosen QD shape and array configuration, we have found the geometrical coefficients N_σ, δ_σ needed for the evaluation of the array effective dielectric function. These two coefficients correspond to two different mechanisms responsible for the modification of the gain in the array. The first mechanism is related to diffraction at individual QDs. Although QDs are assumed to be electrically small and, by this reason, the diffraction can be described within the dipole approximation, its role turns out to be essential owing to the resonant nature of the exciton. It is obvious that the diffraction-induced shift does not depend on the exciton dephasing time τ and the volume fraction f_V of QDs. The second mechanism inducing a frequency shift is a collective effect and is defined by electromagnetic interaction between QDs in the ensemble. This contribution depends on lattice type. The combined effect of both mechanisms can easily be obtained from Eq. 19:

$$\Delta\omega_\sigma = \frac{g_0^\sigma}{\varepsilon_h}(N_\sigma + f_V \delta_\sigma) \quad (22)$$

Different from the geometrical coefficients N_σ, δ_σ, the quantities g_0^σ are phenomenological parameters, which must be invoked with consideration of the outside.

Eq. 17 shows that, in the language of crystal optics,[11] the QD composite being considered is effectively a uniaxial dielectric continuum with the z-axis as its preferred axis. The phenomenon of birefringence is characteristic for this medium: Both ordinary and extraordinary planewave propagations can occur in it.[36] The refractive indices of these waves (n_H and n_E, respectively) are given by:[36]

$$n_H = \sqrt{\varepsilon_H}, \quad n_E = \sqrt{\frac{\varepsilon_E \varepsilon_H}{\varepsilon_H + (\varepsilon_E - \varepsilon_H)\cos^2\theta}} \quad (23)$$

where θ is the angle between the z-axis and the propagation direction. The polarization vector of the ordinary wave is normally directed to the plane constituted by the z-axis and the propagation direction. Eq. 23 shows that $n_H = n_E$ when the propagation direction coincides with the z-axis ($\theta = 0$), and $n_E = \sqrt{\varepsilon_E}$ when the propagation occurs in the xy-plane ($\theta = \pi/2$). Substitution of Eqs. 19–21 and then into Eq. 23 allows us to evaluate the anisotropic macroscopic gain for QD arrays for both polarizations:

$$g_{H,E} = \mathrm{Im}(n_{H,E}) \quad (24)$$

Below we discuss the computation of $g_{H,E}$ by Eqs. 18, 23, and 24.

Numerical Estimates and Discussion

The analytics presented above implies composite medium constituted by uniform two-level insertions in the host. Actually, QDs in ensemble are characterized by essential size dispersion and multilevel structure of the exciton spectrum. These factors provide inhomogeneous line broadening, which is beyond the above consideration. As a first step to take the effect into account, we introduce into analytics phenomenological parameters $g_0 = |\hat{\mathbf{g}}_0|$ and τ as collective characteristics of the ensemble. A simple estimate of g_0 using experimental data can be given in the following way. The parameter g_0 is related to the material gain per dot g_{mat} by $g_{mat} \approx k g_0/2\Gamma_{inh}\sqrt{\varepsilon_h}$, where Γ_{inh} is the inhomogeneous broadening width in the measured QD

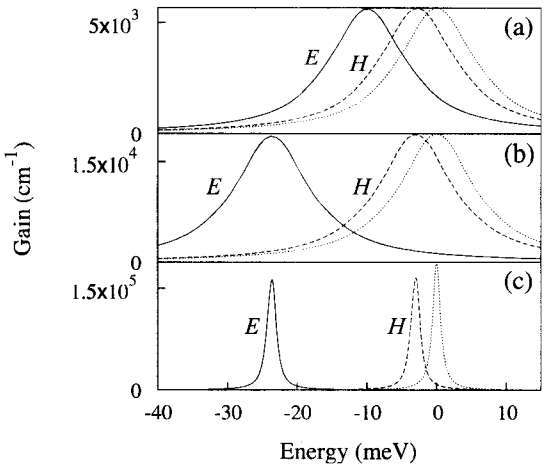


Fig. 3 Macroscopic gain bands for E-polarized and H-polarized fields (solid and dashed lines, respectively) for an array of disk QDs arranged on a cubic (a) and tetragonal (b, c) lattice and different dephasing times $\tau=0.1$ psec (a, b) and $\tau=1$ psec (c). Vertical stacking period $\beta=1$ (a), 0.3 (b, c); QD's volume fraction $f_V=0.02$ (a), 0.07 (b, c); $b/a=0.33$; $a/d=0.25$; host dielectric constant $\varepsilon_h=12.25$; and wavelength $\lambda=1$ μm. Also shown by dotted lines are the gain bands of QD ensembles defined by Eq. 6 with isotropic $\hat{\mathbf{g}}_0$. (From Ref. [5].)

sample. The value of g_{mat} is extracted from experimental measurements of the QD ensemble gain. Letting $g_{mat}\cong 10^5$ cm^{-1} and $\Gamma_{inh}\cong 10^{13}$ sec^{-1},[1] for $\lambda=1300$ nm, one can obtain $|g_0|\cong 10^{14}$ sec^{-1}. The value of the experimentally estimated maximum material gain correlates with theoretical predictions.[37,38] The above estimate is used further for calculations.

Fig. 3 demonstrates the polarization-dependent shift of the resonance energy of macroscopic gain for the case $g_0^E=g_0^H$ at different magnitudes of the effective dephasing time and $\theta=\pi/2$. In the case of disklike QDs, the gain band for H-polarized field directed toward the long axis of the disc possesses the smaller shift. The polarization dependence of the shift results in an appreciable polarization splitting of the gain bands for E-polarization and H-polarization, which become completely resolved for sufficiently large dephasing times (Fig. 3c). Thus radiation from an active composite must be linearly polarized at a given frequency.

As it has been pointed out above, the analytical treatment illustrated by Fig. 3 is invalid at $\beta\ll 1$ and $\beta\gg 1$. Fig. 4 demonstrates polarization splitting of the gain band for cubic (Fig. 4a) and tetragonal lattices (Fig. 4b and c) of disklike QDs, obtained via direct numerical simulation by the minimal autonomic block (MAB) method.[5] Comparison with Fig. 3 shows that for cubic lattices (Figs. 3a and 4a), the polarization splitting effect manifests itself in the same manner, distinguished only in that the peak gains for E-polarized and H-polarized waves are slightly different in the last case. The situation is changed as β decreases (i.e., disks approach each other along the z-axis as far as they are in contact). Instead of the splitting increase demonstrated by Fig. 3b, we see a reverse tendency (Fig. 4b) up to the interchange of the E-polarized and H-polarized peaks for closely stacked pyramids (Fig. 4c). Note that the relative location of the E-polarized and H-polarized peaks in last case correlates with analytical results for quantum wires presented in Ref. [4]. Thus from an electrodynamic point of view, closely stacked QDs form a corrugated quantum wire. The corrugation is responsible for the quantitative difference in the polarization splitting for closely stacked QDs (Fig. 4b) and cylindrical quantum wires:[4] Numerical results for slablike QDs at $\beta\to 0$ correlate well with the analytics. The fact that the tendency inherent in the polarization splitting for closely stacked QDs cannot be derived from analytical treatment reflects both the restrictiveness of the Maxwell Garnet formalism of the effective medium approach utilized in our paper and the use of approximate expressions (Eq. 21) for the geometrical coefficients δ_σ.

Fig. 5 demonstrates the formation of the gain band fine structure in tetragonal lattices of disklike QDs for the E-polarized field (extraordinary wave; Eq. 23) at oblique

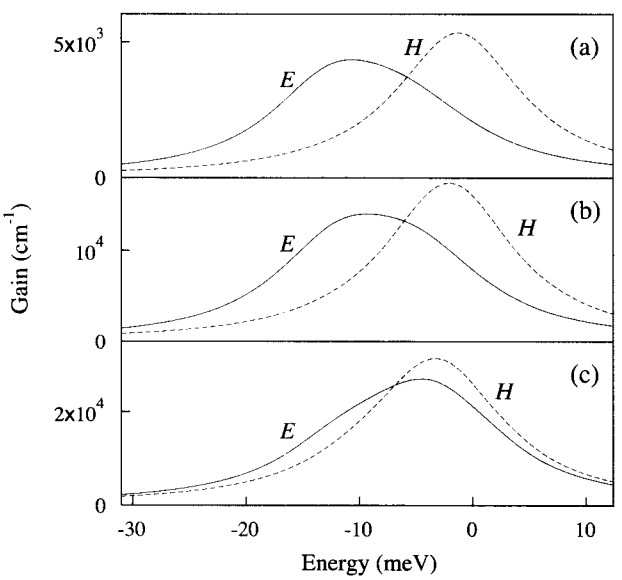


Fig. 4 Macroscopic gain bands for E-polarized and H-polarized fields (solid and dashed lines, respectively) for an array of disklike QDs arranged on a cubic (a) and tetragonal (b, c) lattice. The input parameters ε_h, λ, b/a, and a/d are the same as in Fig. 3. Dephasing time is $\tau=0.1$ psec. Vertical stacking period $\beta=1$ (a), 0.3 (b), and 0.2 (c). QD's volume fraction $f_V=0.02$ (a), 0.07 (b), and 0.1 (c). (From Ref. [5].)

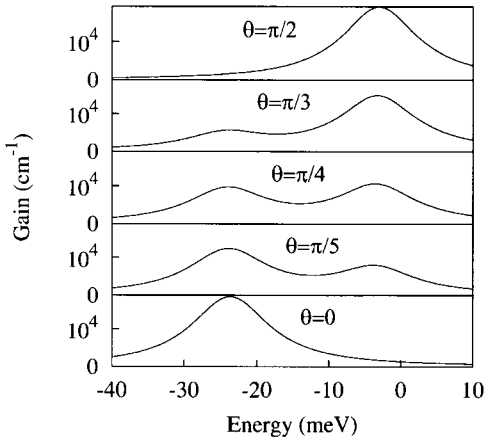


Fig. 5 Macroscopic gain of E-polarized field by disk DQs arranged on a tetragonal lattice for different θ from $\pi/2$ to 0 as indicated. Lattice and individual QD parameters are as in Fig. 3. (b). (From Ref. [5].)

propagation with respect to the z-axis. As is predicted by Eq. 23, the H-polarized field exhibits single resonance coinciding with that for E-polarization at $\theta=0$. Owing to the high symmetry of pyramids, the MAB method simulation for arrays of disklike and pyramidal QDs shows that both models of the QD shape manifest qualitatively similar behavior differing in the shift values and peak amplitudes. It should be noted that the band fine structure illustrated in the particular case given by Fig. 5 must be involved in considering the interpretation of any polarization effects in QD arrays.

In general, also the anisotropic shape of QD causes different oscillator strengths for transitions involving E-polarized and H-polarized fields.[37] To estimate the influence of different gains for different polarizations, we considered the case when the gain for E-polarization is lower and the case when the gain for H-polarization is higher (Fig. 6) compared with the case of equal gains in Fig. 3a and b (parameter g_0 in Eq. 8 is the diagonal tensor). To make the effect clearer, we select also a longer exciton dephasing time. The remarkable result is that the gain peak of E-polarization locates closer to the QD resonance energy because the decrease of the gain peak dominates over the corresponding depolarization coefficient and results in smaller shifts than follow from Eq. 9.

To verify the predicted effects experimentally, a structure with a dense array of QDs was fabricated at the Abram Ioffe Physical Technical Institute (St. Petersburg). The structure represented 12 stacks of arrays of ZnCdSe disklike QDs formed in a self-organized way. The growth, structural, and basic optical properties of the structure are given in Ref. [39]. The areal density of QDs in a single sheet was about 10^{12} cm^{-2}; the ZnCdSe dots had a lateral size of $\sim$4 nm and a height of $\sim$1.2 nm. The dots were incorporated in a ZnSe matrix and separated with 5-nm ZnSe spaces.[39,40] At low excitation density in edge geometry, QD photoluminescence (PL) is linearly polarized[5,40] throughout the PL band with a degree of polarization of 50% of the H-component (alignment as in Fig. 1). This effect is related to the anisotropy in optical transition matrix elements for H-polarizations and E-polarizations originating from the disklike shape of the QDs and the heavy-hole-like nature of the QD exciton. The light-hole-like exciton QD state is also seen in the PL excitation spectrum and is shifted significantly away from the spectral range of interest. With increase in excitation density, gain develops[39] and causes a strong superlinear growth of the PL intensity. Spectra of the H-components and E-components of stimulated emission are shown in Fig. 7b. The predominantly H-polarized component is shifted from the QD resonance revealed in the PL excitation and optical reflectance spectra (Fig. 7a). The degree of polarization of this component strongly increases with excitation density. At the same time, we distinctly observe an appearance of a second separate peak in the E-component at energies closer to the QD resonance. The intensity of the E-polarized peak is much smaller, pointing to a smaller gain coefficient as compared with that for the H-polarized peak at lower photon energy. Therefore the peak location of the E-component corresponds to the predictions of the calculations for the case

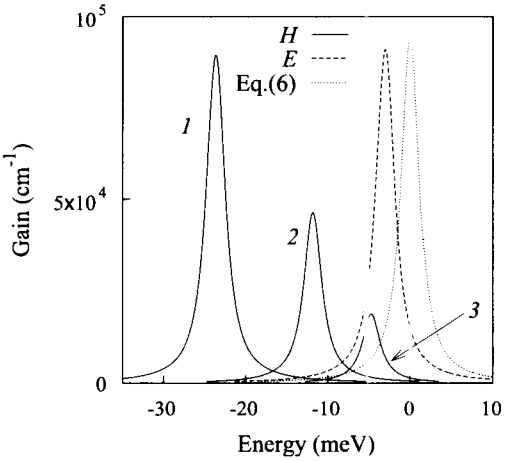


Fig. 6 Macroscopic gain of H-polarized field (solid curves) by disklike QDs arranged on a tetragonal lattice for different anisotropic gains with g^E/g^H = 1.0(1), 4.0(2), 12.0(3), normal incidence $\theta = \pi/2$, τ = 0.5 psec, and other parameters as in Fig. 5. Dashed curve = a gain band for E-polarization; dotted curve = a gain band of a QD ensemble defined by Eq. 6 with isotropic $\hat{\mathbf{g}}_0$. (From Ref. [5].)

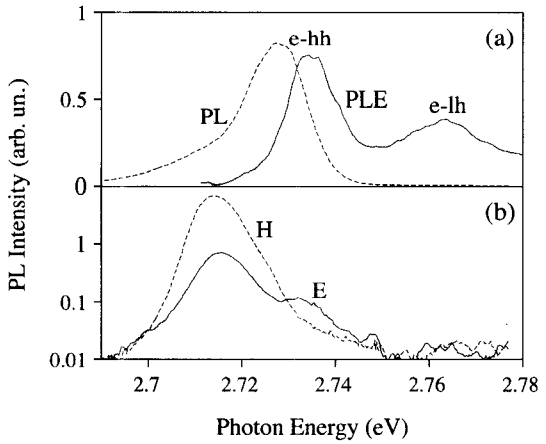


Fig. 7 Photoluminescence (PL) and photoluminescence excitation (PLE) spectra (a) and *E*-polarized and *H*-polarized spectra of stimulated emission in edge geometry (b) for the observation temperature = 7 K, excitation density = 1 MW/cm^2, and energy of exciting photon E_{ex} = 2.88 eV. (From Ref. [4].)

of different gain coefficients for *E*-polarizations and *H*-polarizations (Fig. 6).

Reflectance of Planar Array of Quantum Dots

As was pointed out above, a planar layer comprising a 2-D ensemble of QDs can be treated as an effective QW. As a result, the formalism of investigation of QWs can be extended to QD arrays by introducing effective integral parameters of the array defined by Eq. 4. In particular, it can easily be shown that EBCs (Eq. 3) describe a QW of the thickness L_{QW} with the tensorial dielectric function given by Eq. 17:

$$\varepsilon_H(\omega) = \varepsilon_h - \frac{4\pi i}{\omega}\sigma_{xx}L_{QW}, \quad \varepsilon_E(\omega) = \frac{\varepsilon_h}{1 + 2\xi/L_{QW}}$$

Reflection coefficients for such a QW are given by:[117]

$$r_H(\omega) = \frac{ik\eta_H}{2\sqrt{\varepsilon_h}\cos\theta - ik\eta_H},$$

$$r_E(\omega) = \frac{ik[\eta_H\cos^2\theta + \varepsilon_h^2\eta_E\sin^2\theta]}{2\sqrt{\varepsilon_h}\cos\theta - ik[\eta_H\cos^2\theta - \varepsilon_h^2\eta_E\sin^2\theta]} \quad (25)$$

where θ is the angle of incidence and:

$$\eta_H = L_{QW}[\varepsilon_H(\omega) - \varepsilon_h], \quad \eta_E = L_{QW}\left[\frac{1}{\varepsilon_E(\omega)} - \frac{1}{\varepsilon_h}\right]$$

For spherical particles, after substituting $L_{QW} \to 2R$,[7] Eq. 25 describes the reflection from the planar array of QDs. Thus we state the mathematical equivalence of optical properties of a 2-D periodical layer of QDs and an isolated quantum well. It should be stressed that the mechanisms of electron transport processes and oscillator strengths in each case are essentially different. Nevertheless, the equivalence makes it possible to extend to QD-based planar structures with more complicated configurations (finite-sized QD layer, QD layer in microcavity, several QD layers, etc.) the well-developed mathematical formalism of investigation of quantum wells.[41] Namely this equivalence provides promising potentiality of the derived EBCs for particular electrodynamic problems in QD-based structures. In particular, the threshold current for QD-based lasers can be evaluated by analogy with solutions of corresponding problems for the QW lasers; the EBC method allows us to analyze electromagnetic responses of a QD layer (or a multilayer structure) placed in microcavity—this is very important for the design of QD-based semiconductor lasers.[1]

CONCLUSION

In our paper, we have discussed the electromagnetic response properties of 2-D and 3-D inhomogeneous semiconductor heterostructures—excitonic composites—constituted by QDs imbedded in a transparent host medium. The Maxwell Garnett approach of the effective medium theory and its modification—impedance boundary conditions—have been applied to study 3-D and 2-D structures, respectively. The depolarization shift of the QD gain band and, in anisotropically shaped QDs, the polarization-dependent splitting of this band have been predicted to exist as manifestation of the local field effects. Both types of the exciton localization—strong and weak confinement regimes—with a pronounced nonlocality in the latter regime have been addressed. The basic result is that in both cases, spatially local constitutive relations describe QD composites although the electromagnetic response of an isolated QD in the weak confinement regime is spatially nonlocal. The nonlocality manifests itself in specific dependence of the effective constitutive parameters on the geometry and electronic properties of a single QD.

In our consideration, we have restricted ourselves to the simplest model of excitonic composites—the perfect lattice of identical QDs—whereas experimentally, available structures show intrinsic size dispersion and periodicity violation. These effects may be characterized by a distribution function and taken into account by a theory of irregular excitonic composites. Elaboration of such a theory is the topical question. Another essential restriction of the analysis presented is that nonclassical light–matter interactions have been left beyond the analysis. The quantum

nature of light interacting with an excitonic composite is expected to be of importance for many problems where collections of excitons are involved: quantum computing, electromagnetic fluctuations, etc. Both topical questions will be addressed elsewhere in the future.

ACKNOWLEDGMENT

The research is partially supported through the NATO Science for Peace Program under project SfP-972614.

REFERENCES

1. Bimberg, D.; Grundmann, M.; Ledentsov, N.N. *Quantum Dot Heterostructures*; John Wiley and Sons: Chichester, 1999.
2. *Selected Papers on Linear Optical Composite Materials*; Lakhtakia, A., Ed.; SPIE Optical Engineering Press: Bellingham, WA, 1996.
3. Cho, K. Mechanisms for LT splitting of polarization waves: A link between electron–hole exchange interaction and depolarization shift. J. Phys. Soc. Jpn. **1999**, *68* (2), 683–691.
4. Slepyan, G.Ya.; Maksimenko, S.A.; Kalosha, V.P.; Ledentsov, N.N.; Bimberg, D.; Alferov, Zh.I. Polarization splitting of the gain band in quantum wire and quantum dot arrays. Phys. Rev., B **1999**, *59* (19), 12275–12278.
5. Maksimenko, S.A.; Slepyan, G.Ya.; Kalosha, V.P.; Maly, S.V.; Ledentsov, N.N.; Herrmann, J.; Hoffmann, A.; Bimberg, D.; Alferov, Zh.I. Electromagnetic response of 3D arrays of quantum dots. J. Electron. Mater. **2000**, *29* (5), 494–503.
6. Maksimenko, S.A.; Slepyan, G.Ya.; Ledentsov, N.N.; Kalosha, V.P.; Hoffmann, A.; Bimberg, D. Light confinement in a quantum dot. Semicond. Sci. Technol. **2000**, *15* (5), 491–496.
7. Slepyan, G.Ya.; Maksimenko, S.A.; Kalosha, V.P.; Hoffmann, A.; Bimberg, D. Effective boundary conditions for planar quantum-dot structures. Phys. Rev., B **2001**, *64* (12), no. 125326.
8. Maksimenko, S.A.; Slepyan, G.Ya.; Kalosha, V.P.; Ledentsov, N.N.; Hoffmann, A.; Bimberg, D. Size and shape effects in electromagnetic response of quantum dots and quantum dot arrays. Mater. Sci. Eng., B **2001**, *82* (1–3), 215–217.
9. Slepyan, G.Ya.; Maksimenko, S.A.; Hoffmann, A.; Bimberg, D. Local field effects in an isolated quantum dot: Self-consistent microscopic approach. Phys. Status Solidi, A **2002**, *190* (2), 555–559.
10. Slepyan, G.Ya.; Maksimenko, S.A.; Hoffmann, A.; Bimberg, D. Quantum optics of a quantum dot: Local-field effects. Phys. Rev., A **2002**, *66* (6), no. 063804.
11. Born, M.; Wolf, E. *Principles of Optics*, 6th Ed.; Pergamon Press: Oxford, 1985. Sec. 14.5.
12. Von Faxen, H. Der Zusammenhang zwischen den Maxwellschen Gleichungen für Dielektrika und den atomistischen Ansätzen. Z. Phys. **1920**, *2* (3), 218–229.
13. Khiznjak, N.A. *Integral Equations of Macroscopic Electrodynamics*; Naukova Dumka: Kiev, 1986 (in Russian).
14. Weinstein, L.A. *The Theory of Diffraction and the Factorization Method*; Golem: New York, 1969.
15. Ilyinsky, A.S.; Slepyan, G.Ya.: Slepyan, A.Ya. *Propagation, Scattering and Dissipation of Electromagnetic Waves*; Peter Peregrinus: London, 1993.
16. Hoppe, D.J.; Rahmat-Samii, Y. *Impedance Boundary Conditions in Electromagnetics*; Taylor and Francis: Washington, 1995.
17. Atanasov, R.; Bassani, F.; Agranovich, V.M. Mean-field polariton theory for asymmetric quantum wells. Phys. Rev., B **1994**, *49* (4), 2658–2666.
18. Zaluzny, M.; Nalewajko, C. Coupling of infrared radiation to intersubband transitions in multiple quantum wells: The effective-medium approach. Phys. Rev., B **1999**, *59* (20), 13043–13053.
19. Maksimenko, S.A.; Slepyan, G.Ya. Electrodynamic Properties of Carbon Nanotubes. In *Electromagnetic Fields in Unconventional Structures and Materials*; Singh, O.N., Lakhtakia, A., Eds.; John Wiley and Sons: New York, 2000; 217–255.
20. Maksimenko, S.A.; Slepyan, G.Ya. Electrodynamics of carbon nanotubes. J. Commun. Technol. Electron. **2002**, *47* (3), 235–252.
21. Shubin, V.A.; Sarychev, A.K.; Clerc, J.P.; Shalaev, V.M. Local electric and magnetic fields in semicontinuous metal films: Beyond the quasistatic approximation. Phys. Rev., B **2000**, *62* (16), 11230–11244.
22. Pantell, R.H.; Puthoff, H.E. *Fundamentals of Quantum Electronics*; John Wiley and Sons: New York, 1969.
23. Landau, L.D.; Lifshitz, E.M. *Electrodynamics of Continuous Media*; Pergamon Press: Oxford, 1960.
24. Kleinman, R.E. Low Frequency Electromagnetic Scattering. In *Electromagnetic Scattering*; Uslenghi, P.L.E., Ed.; Academic Press: New York, 1978; 1–28.
25. Kayanuma, Y. Resonant interaction of photons with a random array of quantum dots. J. Phys. Soc. Jpn. **1993**, *62* (1), 346–356.
26. Schmitt-Rink, S.; Miller, D.A.B.; Chemla, D.S. Theory of the linear and nonlinear optical properties

of semiconductor microcrystallites. Phys. Rev., B **1987**, *35* (15), 8113–8125.
27. Hanewinkel, B.; Knorr, A.; Thomas, P.; Koch, S.W. Optical near-field response of semiconductor quantum dots. Phys. Rev., B **1997**, *55* (20), 13715–13725.
28. Keller, O. Local fields in the electrodynamics of mesoscopic media. Phys. Rep. **1996**, *268* (2–3), 85–262.
29. Regelman, D.V.; Mizrahi, U.; Gershoni, D.; Ehrenfreund, E.; Schoenfeld, W.V.; Petroff, P.M. Semiconductor quantum dot: A quantum light source of multicolor photons with tunable statistics. Phys. Rev. Lett. **2001**, *87* (25), no. 257401.
30. Birkedal, D.; Leosson, K.; Hvam, J.M. Long lived coherence in self-assembled quantum dots. Phys. Rev. Lett. **2001**, *87* (22), no. 227401.
31. Borri, P.; Langbein, W.; Schneider, S.; Woggon, U.; Sellin, R.L.; Ouyang, D.; Bimberg, D. Exciton relaxation and dephasing in quantum-dot amplifiers from room to cryogenic temperature. IEEE J. Sel. Top. Quantum Electron. **2002**, *8* (5), 984–991.
32. Verzelen, O.; Ferreira, R.; Bastard, G. Excitonic polarons in semiconductor quantum dots. Phys. Rev. Lett. **2002**, *88* (14), no. 146803.
33. Yu, P.Y.; Cardona, M. *Fundamentals of Semiconductors. Physics and Material Properties*; Springer-Verlag: Berlin, 2001.
34. Fu, Y.; Willander, M.; Ivchenko, E.L. Photonic dispersions of semiconductor-quantum-dot-array-based photonic crystals in primitive and face-centered cubic lattices. Superlattices Microstruct. **2000**, *27* (4), 255–264.
35. Hanamura, E. Very large optical nonlinearity of semiconductor microcrystallites. Phys. Rev., B **1988**, *37* (3), 1273–1279.
36. Yariv, A.; Yeh, P. *Optical Waves in Crystals*; John Wiley and Sons: New York, 1983, Sec. 4.6.
37. Asada, M.; Miyamoto, Y.; Suematsu, Y. Gain and the threshold of three dimensional quantum dot lasers. IEEE J. Quantum Electron. **1986**, *32*, 1915–1933.
38. Asryan, L.V.; Suris, R.A. Inhomogeneous line broadening and the threshold current density of a semiconductor quantum dot laser. Semicond. Sci. Technol. **1996**, *11* (4), 554–567.
39. Strassburg, M.; Kutzer, V.; Pohl, U.W.; Hoffmann, A.; Broser, I.; Ledentsov, N.N.; Bimberg, D.; Rosenauer, A.; Fischer, U.; Gerthsen, D.; Krestnikov, I.L.; Maximov, M.V.; Kop'ev, P.S.; Alferov, Zh.I. Gain studies of (Cd, Zn)Se quantum islands in a ZnSe matrix. Appl. Phys. Lett. **1998**, *72* (8), 942–944.
40. Krestnikov, I.L.; Strassburg, M.; Caesar, M.; Hoffmann, A.; Pohl, U.W.; Bimberg, D.; Ledentsov, N.N.; Kop'ev, P.S.; Alferov, Zh.I.; Litvinov, D.; Rosenauer, A.; Gerthsen, D. Control of the electronic properties of CdSe submonolayer superlattices via vertical correlation of quantum dots. Phys. Rev., B **1999**, *60* (12), 8695–8703.
41. Bastard, G. *Wave Mechanics Applied to Semiconductor Heterostructures*; Les editions de Physique: Les Ulis, 1988.

Quantum Dot Lasers

Mikhail V. Maximov
Ioffe Physico-Technical Institute, St. Petersburg, Russia

Nikolai N. Ledentsov
Technische Universität Berlin, Berlin, Germany

INTRODUCTION

Quantum dot (QD) lasers emerged about decade ago and now they are one of the most rapidly developing and exciting areas in the field of semiconductor optoelectronics. Properly optimized QD lasers show better performance than conventional quantum well (QW) lasers: low threshold current, enhanced temperature stability, high differential efficiency, increased differential gain, and small linewidth enhancement factor resulting in reduced filamentation and chirp. Self-organized QDs offer the possibility to reach emission wavelengths that are larger than those obtainable with quantum wells of the same material system.

In this essay we discuss the fundamentals of QD lasers. The criteria for the optimization of QD lasers for different applications are considered. Experimental demonstrations of the advantages of QD lasers are presented. The latest experimental results on edge-emitting and vertical cavity surface emitting QD lasers are reviewed. Possible short-term commercial applications of QD lasers are also discussed.

ADVANTAGES OF QUANTUM DOT LASERS

The concept of carrier confinement has been of primary importance in the development of semiconductor laser. The invention of the double heterostructure,[1,2] in which carriers in a narrow-gap material are confined by wide bandgap barriers, made possible the first continuous wave (CW) operation of semiconductor lasers at room temperature and their practical implementation. The next breakthrough occurred when Dingle and Henry proposed the idea to "exploit quantum effects in heterostructure semiconductor lasers to produce wavelength tunability" and to achieve "lower lasing thresholds" via "the change in the density of states which results from reducing the number of translational degrees of freedom of the carriers".[3] It was also shown that if the number of translational degrees of freedom of charge carriers is decreased below two, a singularity occurs in the density of states (Fig. 1). This singularity increases light absorption or light amplification (gain). An ultimate case of size quantization is realized in QDs. A QD is a coherent inclusion of a narrow-gap material in a wide-gap matrix in which electrons are quantized in all three spatial directions. Thus a single semiconductor QD exhibits a discrete δ-function-like energy spectrum similar to that in a real atom, keeping the advantage of direct current injection, impossible in other types of lasers based on atomic transitions.

The physical advantages of QD lasers resulting from δ-function-like density of states are:

- Low threshold current
- High material gain
- High temperature stability of the threshold current (characterized by characteristic temperature T_0)
- High differential efficiency
- High differential gain
- Small α-factor and correspondingly reduced filamentation and chirp (shift of the lasing wavelength with current)

In addition, QD medium has some advantages, which are not directly related to size quantization effects:

- The possibility to reach emission wavelengths that are larger than those obtainable with quantum wells of the same material system. For instance, lasing at wavelengths up to 1.32 μm is demonstrated for InAs QDs on GaAs substrate, in contrast to QW devices, in which lasing is limited to approximately 1.15 μm.
- Dramatically reduced carrier lateral transport. Carrier capture in QDs is very fast and carrier localization in QDs is very efficient. This leads to a number of advantages of using QD medium in laser applications: larger catastrophic optical damage threshold due to decreased facet overheating, increased lifetime, enhanced radiation hardness, and the possibility to

N.N. Ledentsov is on the leave from Ioffe Physico-Technical Institute.

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013519

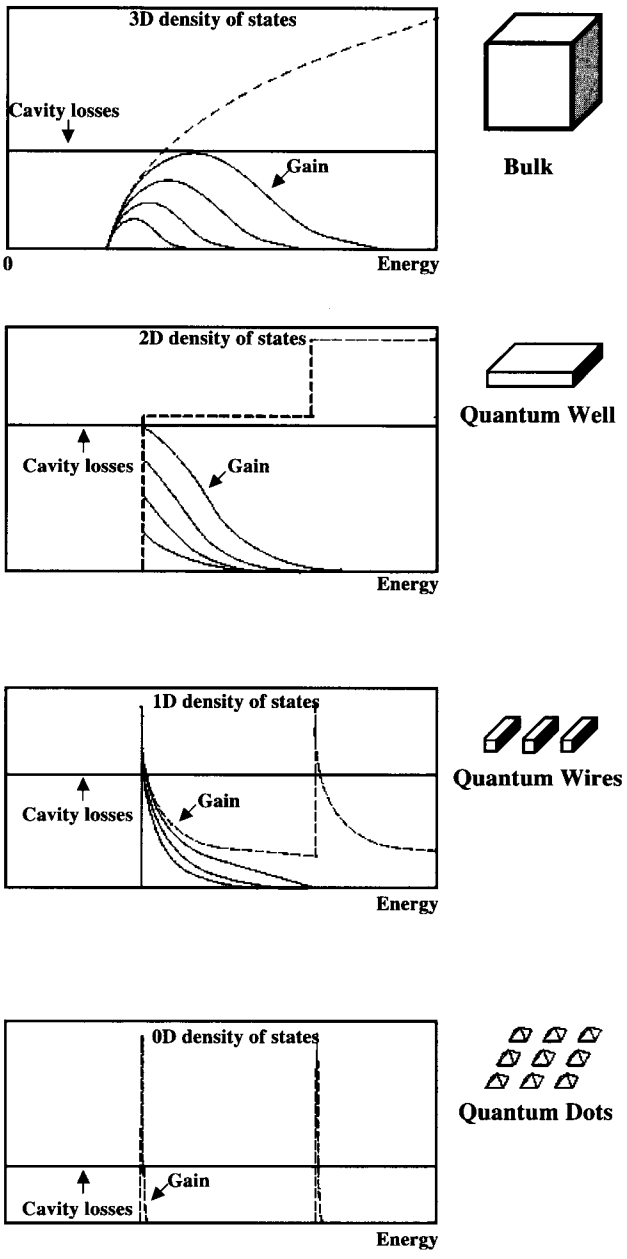


Fig. 1 A schematic sketch of the density of states (dashed curves) and gain spectra (solid curves) in bulk material, quantum well, quantum wires, and quantum dots. Schematic morphology is shown on the right.

fabricate narrow stripes via etching through active area without increase in surface recombination.

Some of the basic advantages of QD lasers such as improvement in the temperature stability of the threshold current,[4] high material gain,[5] and the possibility of remarkable reduction in the laser threshold were already described in the first theoretical works on QD lasers. These works, however, were generally based on simplified assumptions such as infinite barriers, no QD size fluctuations, one confined electron and hole level, and ultrafast energy relaxation of injected carriers. More recent theoretical models had to take into account such complications as finite barriers, many electron and hole levels (effect of excited states), QD size fluctuations, many body effects, radiative and nonradiative recombination in the optical confinement layer (OCL; wide-gap matrix), charge neutrality violation in QDs, etc. It was shown that the characteristics of QD lasers depend dramatically on the parameters characterizing QD array (QD lateral size, height and corresponding position of energy levels, QD density, and size dispersion) as well as on the structure design: thickness of the OCL, doping profiles in the cladding layers and in the OCL, the band offsets at the interface between the OCL and the cladding layers. Depending on the abovementioned parameters the performance of QD lasers can be very good or poor. To get the best performance, an optimization of QD array parameters should be done consistently with the optimization of the structure design. Optimization for certain applications (low threshold current, high output power, etc.) should be done according to different criteria, and in many applications a combination of these criteria needs to be taken into account.

High temperature stability of the threshold current as the main fingerprint of QD lasers was demonstrated already in the first injection laser based on self-organized QDs.[6] Further realization of the advantages of QD lasers was associated with the improvement in QD growth and the development of more sophisticated QD growth techniques. At present, all the advantages of QD lasers are basically experimentally proved. At the same time we believe that there is still room for improvement in the characteristics of QD lasers. Experimental and theoretical studies of QD lasers are exploding in scientific areas and further progress in the field of QD lasers can be foreseen. Recently, the first commercial QD lasers have been announced.[7]

QUANTUM DOTS GROWTH

Realization of QD lasers became possible owing to the progress in self-organized growth. The deposition of a material with a lattice constant different from that of the substrate was shown to result in a spontaneous formation of three-dimensional (3-D) islands (for a review, see Ref. [8]). In most practical cases, for instance for the deposition of InGaAs on GaAs, first a 2-D layer grows on the substrate and island formation occurs only after a

certain critical thickness is exceeded. Thus QD array is situated on top of a thin initial planar layer (so-called wetting layer). This growth mode is referred to as Stranski–Krastanow growth mode. Under certain growth parameters, arrays of highly uniform coherent islands can be formed. Each island generates long-range strain fields in the substrate and interaction of the island via these fields favors island lateral ordering in a 2-D lattice.

The characteristics of QD lasers depend dramatically on the parameters characterizing QD array.[9] Already in the first injection QD laser based on QDs formed in simple Stranski–Krastanow growth mode, insufficient carrier localization in QDs and their thermal escape to the wetting layer and the OCL were shown to be a disadvantage, deteriorating the device performance at high temperatures.[6] As simple Stranski–Krastanow growth provides only limited control of the interrelated size and density of the islands, extensive investigations have been conducted to overcome these limitations and to optimize the properties of QD arrays. Novel advanced approaches such as vertically coupled QDs[10] lead to a considerable improvement in the performance of QD lasers.

Present best performance QD lasers are based on QDs formed by the technological approach that we refer to as activated alloy phase separation (AAPS).[11,12] In this approach the original Stranski–Krastanow InAs islands are formed by depositing a certain amount (D_{IS}) of InAs (D_{IS}=1.7–3 ML). The original islands are then overgrown with an average thickness H of $In_xGa_{1-x}As$ alloy layer (H=0–6 nm, x=0–0.20). The strain field induced by each island leads to a migration of In atoms toward it and, correspondingly, to an increase in its effective volume. The increase in the original island volume as well as partial strain relaxation in such QDs embedded in an $In_xGa_{1-x}As$ layer and a narrowing of the matrix bandgap results in a higher confinement of electron and hole levels with a corresponding red-shift of the photoluminescence (PL) emission.[12,13] Lasers based on QDs formed by the AAPS enable it to reach lasing wavelengths above 1.3 μm for the structures grown on GaAs substrates, which is not possible with the use of InGaAs QWs. The AASP process sensitively depends on the size and density of the InAs nanostressors, i.e., the InAs deposition amount (D_{IS}), as well as on the thickness (H), and on the composition (x) of the $In_xGa_{1-x}As$ alloy layer. To achieve high QD confinement energy (long emission wavelength), narrow QD size distribution, and simultaneously prevent formation of dislocations and large dislocated clusters, a careful optimization of growth parameters is required and described in detail in Ref. [12].

A growth sequence very similar to AAPS is referred to by the authors as a "dots in a well" (DWELL) design.[14] In the DWELL design the original Stranski–Krastanow islands are formed on a several-nanometer-thick $In_xGa_{1-x}As$ buffer layer. To the best of our knowledge QDs formed by AAPS and DWELL have very similar properties. More details on the growth and optical properties of InAs–GaAs QDs emitting in long-wavelength optical region can be found in review.[15]

EQUILIBRIUM VS. NONEQUILIBRIUM CARRIER DISTRIBUTION IN QUANTUM DOT ARRAY

Depending on the sample temperature and the confinement of electron and hole energy levels with respect to

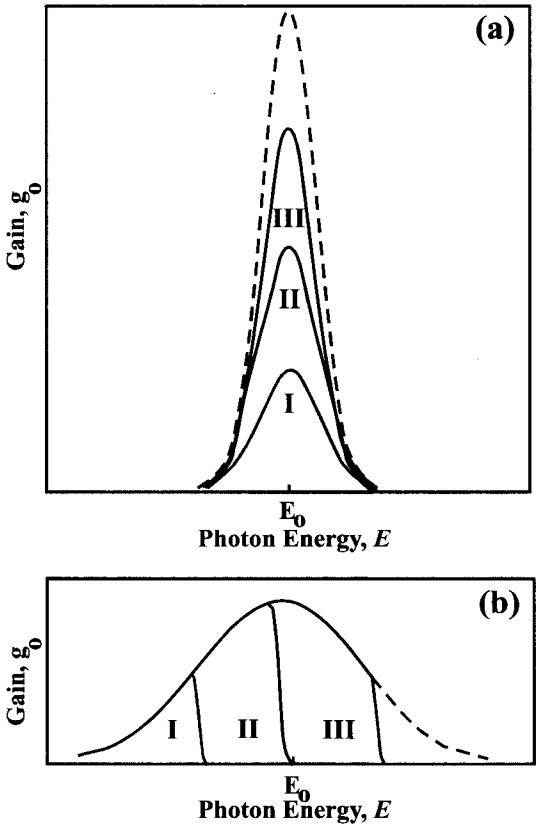


Fig. 2 Gain spectra for the cases corresponding to nonequilibrium (a) and equilibrium (b) filling of QDs. The curves are numbered in ascending order of population inversion in a QD. The dashed curve corresponds to the population inversion equal to unity. E_0 is the energy of the ground state transition in an average sized QD. (From Ref. [9].)

the continuum, the carrier distribution within the QD array can be either equilibrium or nonequilibrium.[9] At high temperatures or (and) low confinement energies the characteristic times of thermally excited escapes of electrons and holes from a QD are small compared to the radiative lifetime in QDs. Thermal carrier emission, lateral transport via the wetting layer and matrix, and then recapture result in the quasi-equilibrium distribution of carriers in the array with the corresponding quasi-Fermi levels determined by the pumping level. As a consequence of such carrier redistribution, the level occupancies (and numbers of carriers) in various QDs will differ.

At low temperatures or (and) efficient confinement of electron and hole energy levels, the radiative lifetime in QDs is small compared with the characteristic times of thermally excited escapes of the carriers from a QD. A carrier captured by a QD having no time to leave a QD recombines in the same QD. The redistribution of carriers from one QD to another and establishment of quasi-Fermi levels for the conduction and valence bands do not occur. As the initial numbers of carriers injected into various QDs are the same, the level occupancies are also the same there.

Carrier distribution dramatically affects the dependence of gain spectra on the injection current.[9] In case of nonequilibrium distribution, the gain spectrum copies the shape of the curve corresponding to the dispersion of QD ground-state transition energies (Gaussian). The spectra are scaled along the vertical axis by a factor equal to the mean population inversion in QDs (Fig. 2a). In case of quasi-equilibrium the "gradual" filling of the curve occurs with the increase in pumping level (Fig. 2b).

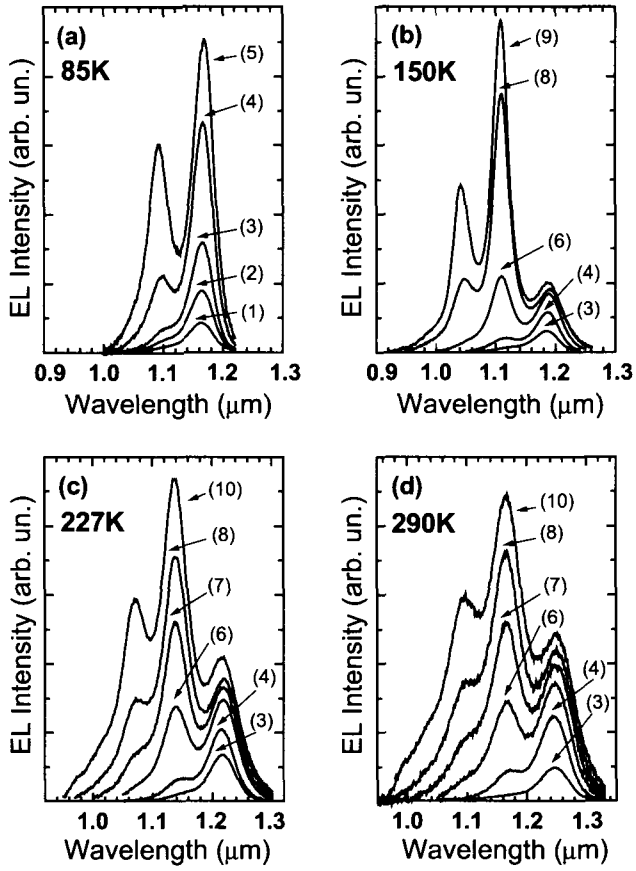


Fig. 3 Electroluminescence spectra from a 200-μm-long stripe at different injection currents and temperatures: (1) 12.5, (2) 25, (3) 50, (4) 125, (5) 250, (6) 500, (7) 1250, (8) 2500, (9) 3750, and (10) 5000 A/cm^2.

SPONTANEOUS EMISSION AND GAIN

The first theoretical simulations of QD lasers were based on the assumption of one electron and one hole level in a QD. In actual QDs the energy level structure is much more complex.[13,16] In a pyramidal QD with base length larger than 14 nm there are several electron and hole levels.[16] Because of piezoelectric effect and Coulomb interaction there is no degeneracy, even accidental, of the excited states. However, several excited-state transitions may act as one "effective" transition if their energies are close. The absorption (gain) spectrum of the "effective" transition will be a superposition of absorption (gain) spectra corresponding to individual excited-state transitions with close energies. The maximal absorption (gain) of this superposition can be higher than that of the ground-state transition.[17]

Typical electroluminescence (EL) spectra of a laser based on QDs with narrow size distribution are shown in Fig. 3. A 200-μm-long stripe was used. Short stripes are essential to measuring the EL spectrum shape correctly. The longer the stripe the higher the probability of high energy photons corresponding to the emission from excited states to be reabsorbed and contribute to the final emission from lower energy levels. By contrast, lower energy photons corresponding to the emission from the ground state easily leave the structure. The reabsorption effect can modify the EL spectra. Moreover, in short stripes, the spontaneous emission can be studied even at high current densities. At the lowest excitation density, the emission mainly originates from the ground-state transition. At higher excitation densities, the first and the second short wavelength peaks become apparent. Strong level filling occurs at high injection currents because of the limited number of QD energy states. At all temperatures, EL lines from the ground- and excited-state transitions are relatively narrow (40–60 nm) and well re-

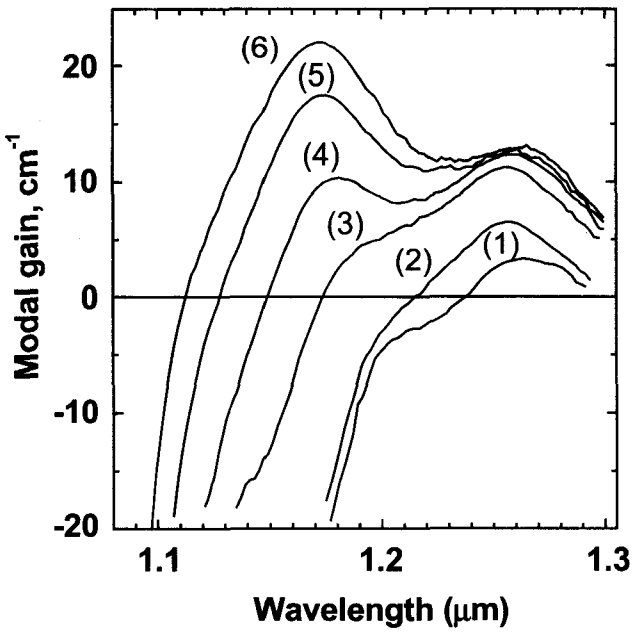


Fig. 4 Modal gain spectra at different injection currents: (1) 98, (2) 195, (3) 487, (4) 975, (5) 1950, and (6) 2925 A/cm^2.

solved. Emission due to the wetting layer/InGaAs layer is also seen in the spectra at 0.97 μm at room temperature (the QDs are embedded in a QW formed by the wetting layer and InGaAs layer with nominal thickness H). Each of the two short wavelength QD peaks is likely to be a superposition of two or more closely located peaks from transitions involving different electron and hole sublevels. At 80 K, the lasing starts at a current density of 500 A/cm^2, whereas at 290 K the device is in the spontaneous emission mode up to 5 kA/cm^2.

Modal gain spectra of a laser structure based on three sheets of QDs formed by the AAPS are shown in Fig. 4. The surface density of QDs estimated from TEM images was 4–5 $\times 10^{10}$ cm^{-2}. The spectra are measured by the variable-stripe length method with varying current injection. The maxima at 1.254 and 1.18 μm correspond to the ground- and first excited-state transitions. At current densities of about 900 A/cm^2 the gain of the ground-state transition saturates. The gain saturation effect is a consequence of the limited surface density of QDs, inhomogeneous broadening of the QD array, and the limited QD radiative emission rate. The maximum modal gain for the transition from the ith electron level to the jth hole level is:[17]

$$g_{ij}^{\max} = \frac{\xi}{4}\left[\frac{\overline{\lambda_{ij}}}{\sqrt{\overline{\epsilon_{ij}}}}\right]^2 d_{ij}\frac{1}{\tau_{ij}^{QD}}\frac{\hbar}{(\Delta\varepsilon)_{\text{inhom},ij}}\frac{\overline{\Gamma}_{ij}}{\bar{a}}N_S Z_L \qquad (1)$$

where

ξ is a numerical constant ($\xi = 1/\pi$ and $\xi = 1/\sqrt{2\pi}$ for Lorentzian and Gaussian QD-size distribution functions, respectively);
$\overline{\lambda_{ij}} = 2\pi\hbar c/\overline{E_{ij}}$ (a wavelength at transition $\overline{E_{ij}}$;
d_{ij} is the degeneracy of the transition;
$\bar{a}$ is the mean size of QDs;
Z_L is the number of QD layers;
N_S is the surface density of QDs in a layer;
$\overline{\epsilon_{ij}}$ is a dielectric constant;
τ_{ij}^{QD} is the spontaneous radiative lifetime for the $i \rightarrow j$ transition;
$(\Delta\varepsilon)_{\text{inhom},ij}$ is the inhomogeneous line broadening for the $i \rightarrow j$ transition;
$\overline{\Gamma}_{ij}$ is the modal optical confinement factor in a QD layer (along the transverse direction in the waveguide) for the $i \rightarrow j$ transition.

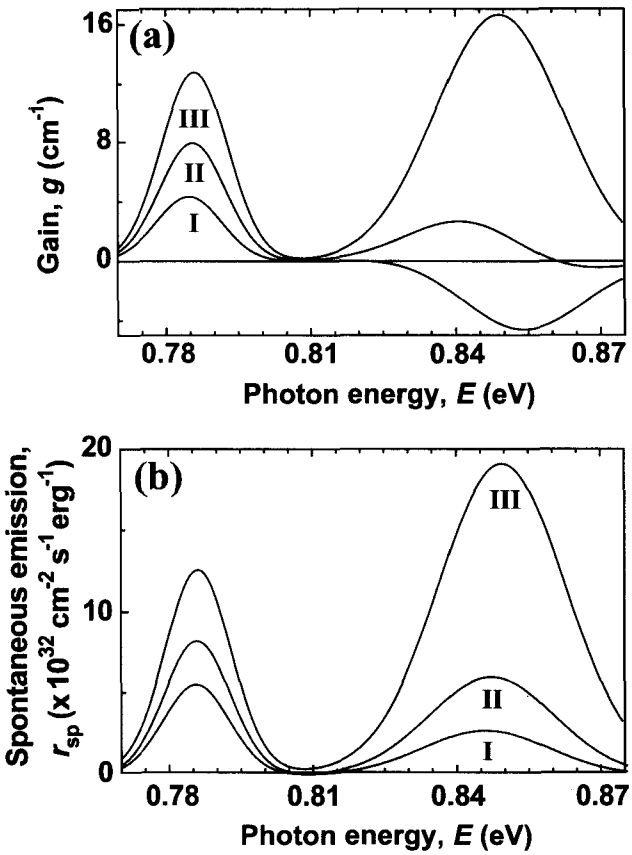


Fig. 5 Gain (a) and spontaneous emission spectra (b) at different injection currents for a model structure in case of large separation of the mean transition energies and small inhomogeneous broadening. I: j=6.8 A cm^{-2}, II: j=18 A cm^{-2}, $j_3=\infty$ (saturated g and r_{sp}). (From Ref. [17].)

The measured maximum (saturated) modal gain of the ground-state transition is 12 cm^{-3} (4 cm^{-1} per QD sheet). For a laser structure based on 10 stacks of similar QDs the measured maximum modal gain was above 25 cm^{-1}.[11]

In equilibrium filling of QDs, the general relationship between the gain and spontaneous emission spectra has been shown to hold for an arbitrary number of radiative transitions and an arbitrary QD-size distribution.[17] Thus gain spectrum can be derived from the measured spontaneous emission spectrum (and vice versa).

EFFECT OF EXCITED-STATE TRANSITIONS ON THE THRESHOLD CHARACTERISTICS OF A QUANTUM DOT LASER

Depending on the relationship between the separation of the ground- and excited-state transition energies and the inhomogeneous broadening, either a smooth or a step-like change in the lasing wavelength with losses occurs.[18,19] The multiple transition effect on threshold characteristics, in particular on the lasing wavelength, was theoretically studied in Ref. [17]. We consider two optical transitions and assume that maximal gain for the "effective" excited-state transition is higher than that for the grown-state transition. The following situations can arise:[17]

1. Large separation of the mean transition energies and small inhomogeneous broadening. The gain and the spontaneous emission spectra for different injection current levels are shown in Fig. 5a and b, respectively. As the injection current density at the lasing threshold is a monotonically increasing function of the losses β, the dependence of the laser characteristics on the injection current density will also represent the dependence on the losses, and vice versa. At $\beta<g_1^{max}$, the gain peak is near the energy of the ground-state transition ($\overline{E_1}$) and, hence, lasing occurs via this transition ($\lambda \approx \overline{\lambda_1}$, where λ is the lasing wavelength and $\overline{\lambda_1}$ is the wavelength of the ground-state transition). For β, somewhat less than g_1^{max}, there is a shift in the peak position, and for $g_1^{max}<\beta<g_2^{max}$, the peak is seen near the energy of the excited-state transition ($\overline{E_2}$). Lasing then occurs via the excited-state transition ($\lambda \approx \overline{\lambda_2}$, where λ is the lasing wavelength and $\overline{\lambda_2}$ is the wavelength of the exited-state transition). The lasing wavelength and threshold current density j_{th} are shown in Fig. 6 as a function of losses. The break in $\partial j_{th}/\partial\beta$ at $\beta \approx g_1^{max}$ reflects the shift of the lasing wavelength. As the modal gain peak is equal to the losses at lasing threshold, it is obvious that the curve for the threshold current density against losses also represents the gain peak vs. the injection current density (the gain–current curve) if the abscissa and the ordinate are interchanged.
2. Small separation of the mean transition energies and large inhomogeneous broadening. The character of the change in the gain and the spontaneous emission spectra with injection current as well as the lasing wavelength dependence on losses are shown in Figs. 7 and 8, respectively. There is a smooth change in λ with β. The lasing wavelength λ can be significantly

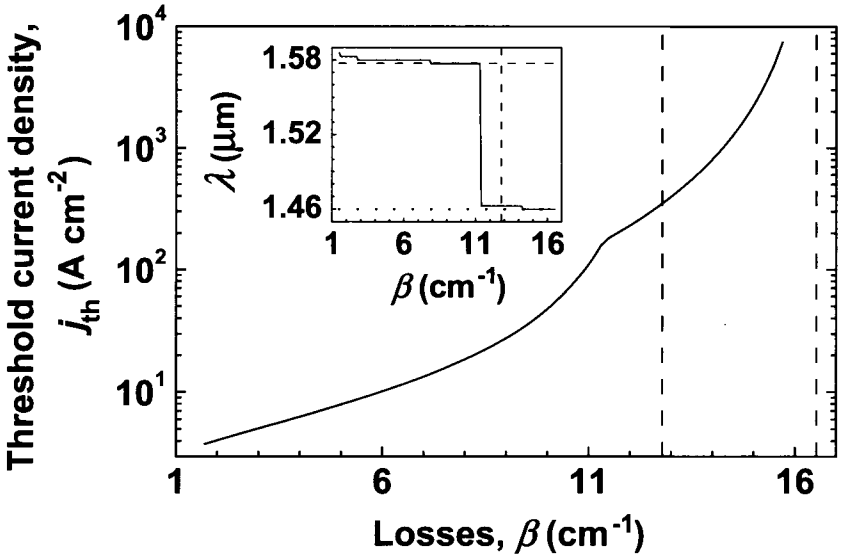


Fig. 6 Threshold current density and lasing wavelength (the inset) against losses. The parameters are the same as in Fig. 5. The vertical dashed lines are for g_1^{max} and g_2^{max}. The horizontal dashed and dotted lines in the inset are for $\overline{\lambda_1}$ and $\overline{\lambda_2}$. (From Ref. [17].)

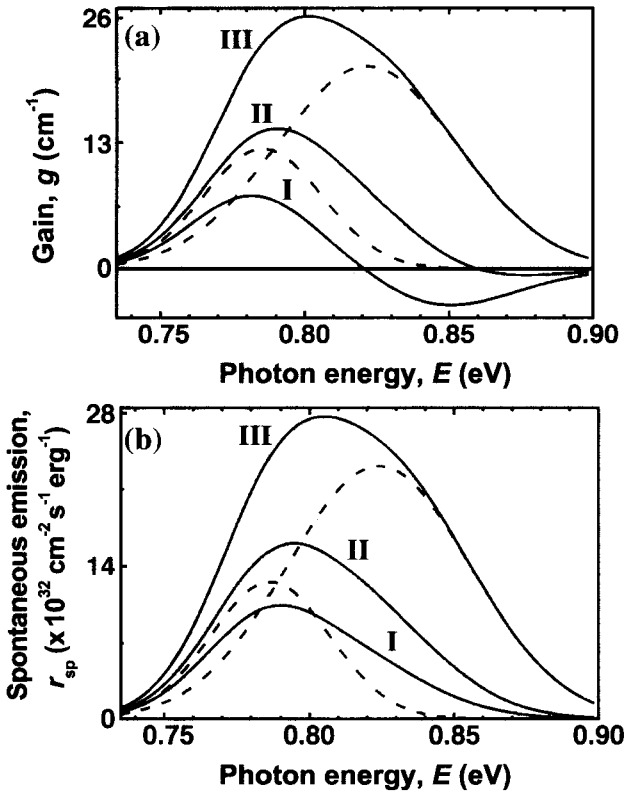


Fig. 7 Gain (a) and spontaneous emission (b) spectra at different injection currents for a model structure in case of small separation of the mean transition energies and large inhomogeneous broadening, I: j=20.7 A cm^{-2}, II: j=39.7 A cm^{-2}, $j_3=\infty$ (saturated g and r_{sp}). (From Ref. [17].)

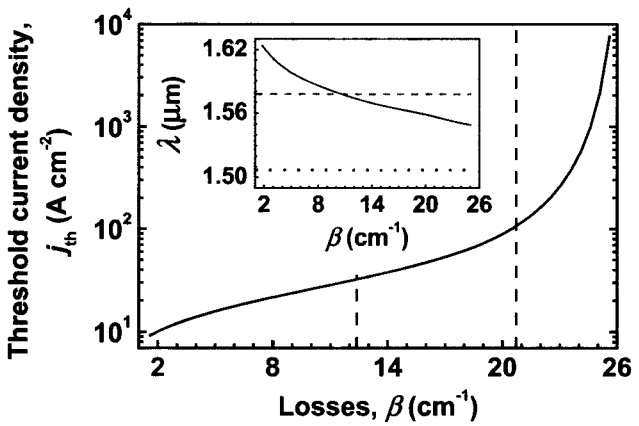


Fig. 8 Threshold current density and lasing wavelength (the inset) against losses. The parameters are the same as in Fig. 7. The vertical dashed lines are for g_1^{max} and g_2^{max}. The horizontal dashed and dotted lines in the inset are for $\overline{\lambda_1}$ and $\overline{\lambda_2}$. (From Ref. [17].)

below $\overline{\lambda_1}$ for high losses. Nevertheless, it is hard to construct a situation where λ approaches $\overline{\lambda_2}$. This is because the position of the gain spectrum peak is determined by the overlap of the transition lines, and the inhomogeneous broadening of the excited-state transition is always larger than that of the ground-state transition. Thus the position of the gain spectrum peak is closer to that of the ground-state transition. Therefore when the lasing wavelength changes smoothly with losses, it is impossible to derive information on the energy of the excited-state transition directly from this dependence.

Near-threshold lasing and EL spectra of a laser structure based on three sheets of QDs formed by the AAPS (the same as was used to measure spontaneous emission and gain spectra; Figs. 3 and 4) are shown in Fig. 9. The inhomogeneous broadening is less than the separation of the mean transition energies (situation 1 takes place). For a long stripe, where the external loss is small, the gain of the ground-state transition is sufficient to overcome the total loss, and lasing occurs via the ground-state

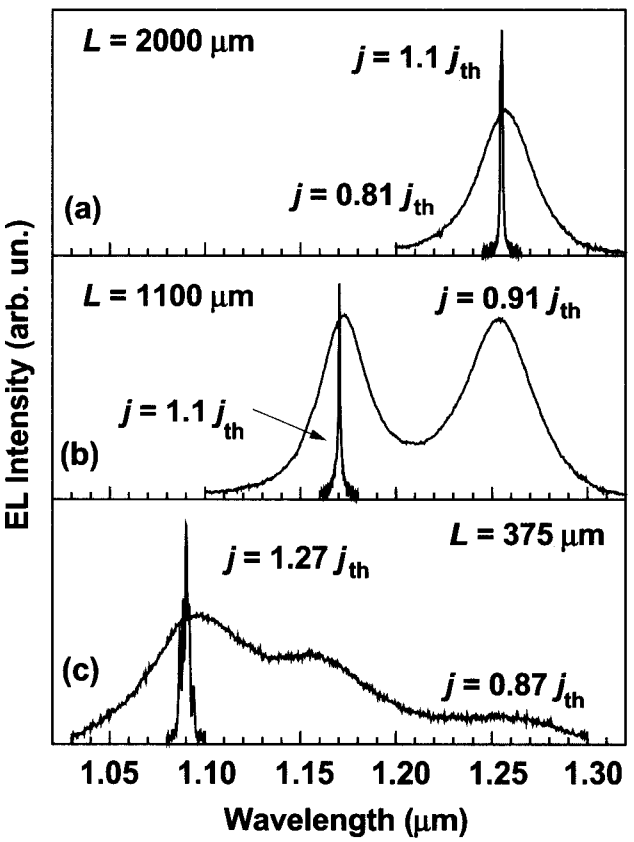


Fig. 9 Electroluminescence spectra showing spontaneous emission and lasing. Different stripe lengths are used.

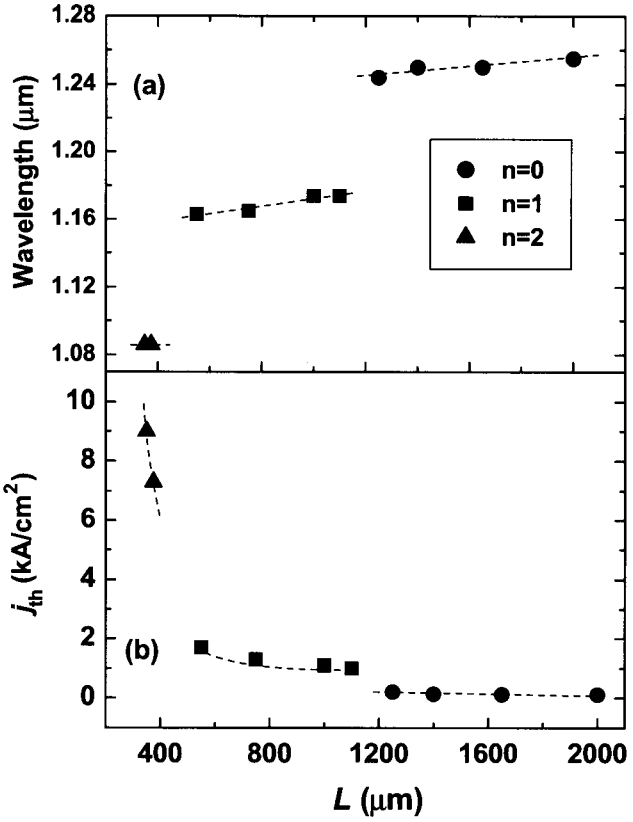


Fig. 10 Dependence of the lasing wavelength (a) and the threshold current density (b) on the cavity length. $n=0$, 1, and 2 denote the ground, first, and second excited states, respectively.

transition (Fig. 9a). In shorter stripes, with larger external loss, the gain of the ground-state transition is not sufficient to overcome the total loss, and the lasing proceeds via the first excited-state transition (Fig. 9b); the threshold current density increases dramatically. The shortest devices lase via the second excited-state transition (Fig. 9c) at a very high threshold current density. The dependence of the lasing wavelength and the threshold current density on the cavity length is summarized in Fig. 10. When the lasing occurs via the ground-state transition (for cavities longer that 1.2 mm) the threshold current density increases only slightly as the cavity length decreases.

THRESHOLD CURRENT DENSITY

Lower threshold current density in QD lasers as compared to those based on quantum wires and QWs was predicted in early theoretical works and greatly motivated experimental and theoretical research in the field of QD lasers. More detailed theoretical studies revealed dramatic dependence of threshold current on the parameters of QD array among which the most important are the QD density and the inhomogeneous broadening.[9] Thus, for instance, it was shown that lasing could never be achieved if QD density is less than certain critical density or inhomogeneous broadening exceeds the critical value. In

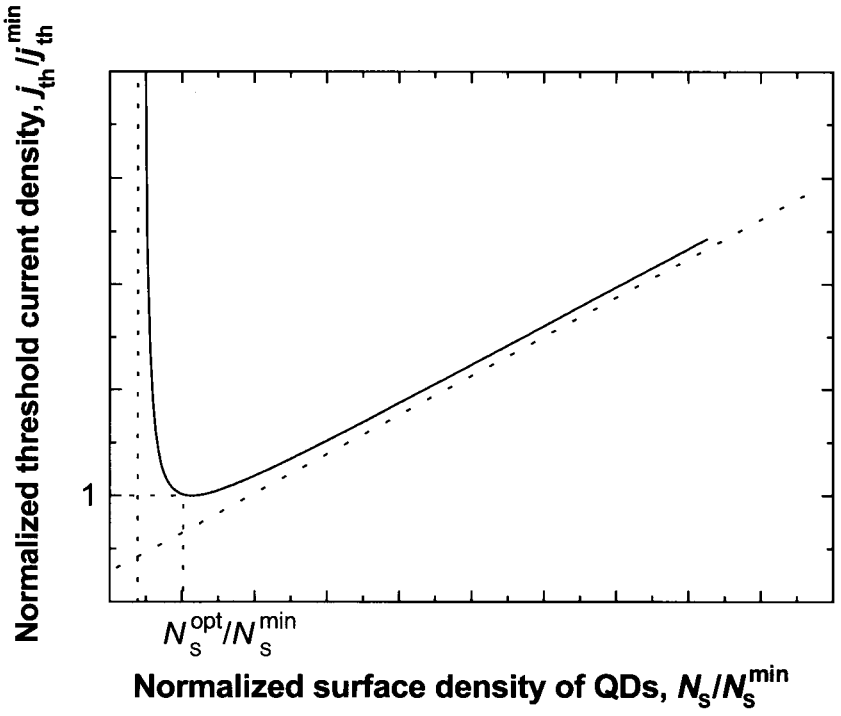


Fig. 11 Normalized threshold current density vs. the normalized surface QD density for a model structure. (From Ref. [9].)

Ref. [20], the following inequality for interrelated tolerable values of structure parameters was obtained:

$$\frac{4}{\xi}\left[\frac{\sqrt{\varepsilon}}{\lambda_0}\right]^2 \tau_{QD}\frac{a}{\Gamma}\frac{(\Delta\varepsilon)_{inhom}}{\hbar}\frac{1}{N_S}\frac{1}{L}\ln\frac{1}{R} \leq 1 \quad (2)$$

where most values are defined in Eq. 1, and where R is the facet reflectivity and L is the cavity length.

Hence Eq. 2 relates the tolerable values of the three parameters, N_S, $(\Delta\varepsilon)_{inhom}$, and L, to each other. The critical tolerable value of each of the three parameters is related to the given values of the other two. The more perfect the QD array [the less the $(\Delta\varepsilon)_{inhom}$] or the longer the cavity the less is the minimum surface density of QDs N_S^{min}. The denser the QD ensemble (the greater the N_S) or the longer the cavity the greater is $(\Delta\varepsilon)_{inhom}^{max}$. The more perfect the QD array or the denser the QD ensemble the less is the minimum cavity length L^{min}.

When one of the structure parameters is close to its critical tolerable value, the mean (averaged over the lateral direction) electron and hole level occupancies in QDs tend to unity—the electron and hole levels are fully occupied. This demands infinitely high free-carrier densities in the OCL and, correspondingly, infinitely high injection current. For the advantages of a QD laser to be attained in practice, structure parameters should be well away from the critical values. Fig. 11 shows the dependence of the normalized threshold current on the normalized surface QD density for a model structure.[9,20] There is an optimal surface QD density minimizing threshold current. For $N_S < N_S^{min}$, the lasing is impossible to attain (the gain cannot overcome losses).

Recent progress in crystal growth has resulted in arrays of In(Ga,As)–GaAs QD with a relatively small (<10%) inhomogeneous broadening. Low threshold current densities of 26 A/cm^2[14] and 19 A/cm^2[21] at 300 K, and 6 A/cm^2 at 4 K[22] for long devices with HR facet coatings based on a single sheet of QDs have been achieved. However, differential efficiency of such lasers was low. In commercial lasers, high differential efficiency is required and, thus, HR coatings cannot be applied to the front facet. Here the measure of quality is the value of threshold

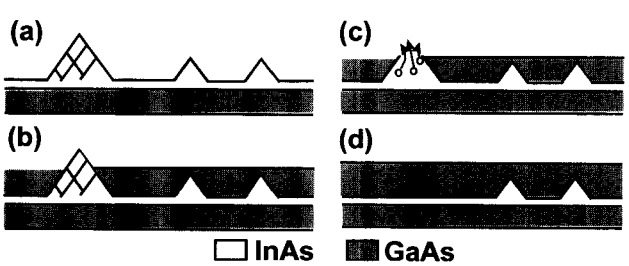


Fig. 12 A schematic illustrating reduction of dislocation density.

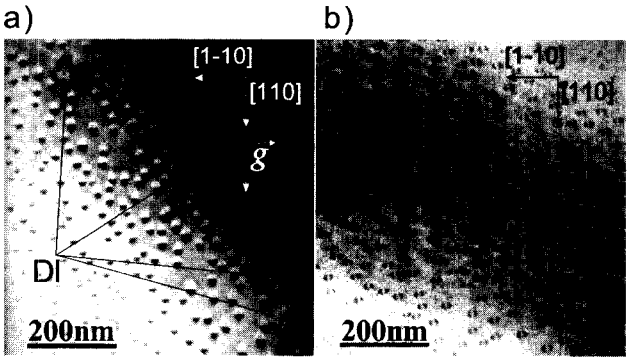


Fig. 13 Plan view bright field (220) TEM images of the structures grown without (a) and with (b) the annealing step. DI denotes dislocated islands.

current density per single QW or QD layer. The best values of 7–10 A/cm^2 per QD layer were realized in devices with a large number of QD stacks[11] and high differential efficiency. A record low transparency current of 6 A/cm^2 per dot layer, an internal quantum efficiency of 98%, and an internal loss below 1.5 cm^{-1} have been demonstrated in a QD laser grown by MOCVD.[23]

DEFECT REDUCTION TECHNIQUES

Growth of QDs especially large-sized ones with high confinement energy for electrons and holes may be accompanied by the formation of defects and large dislocated clusters. This problem can be eliminated by applying specially developed in situ defect-reduction techniques, which permits selective elimination of dislocated objects without affecting coherent QDs.[24] A schematic diagram illustrating the reduction of dislocation density is shown in Fig. 12. The QDs are overgrown with a thin (1 nm) GaAs layer (Fig. 12b) at a temperature corresponding to that of QD formation (480°C). Then the substrate temperature is raised to 600°C. The thin cap layer only partly covers the large dislocated InAs clusters allowing the InAs to evaporate in the annealing step (Fig. 12c). The smaller coherent QDs, being completely covered, are only weakly affected. Fig. 13 shows transmission electron microscopy (TEM) images for the samples grown with and without the annealing step. The density of dislocated islands in the sample grown without the annealing step is $\sim 10^9$ cm^{-2}. No dislocated islands are revealed in the structure grown with the annealing step.

The application of defect reduction techniques is especially important in case of lasers based on multiple QD stacks. Even small density of dislocations and large dislocated clusters in the first QD sheet result in a drastic deterioration of the properties of upper QD sheets. The use

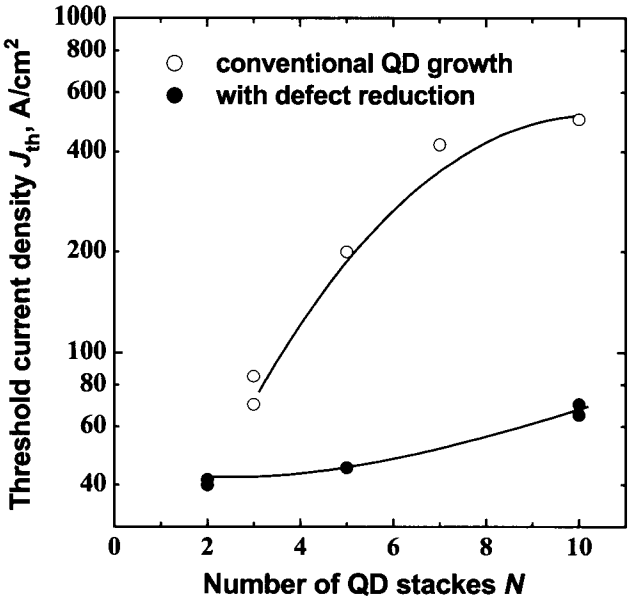


Fig. 14 Threshold current densities for the structures with various numbers of QD stacks grown without and with defect reduction. A design with negligible external losses is used.

of defect reduction techniques results in a dramatic improvement in the characteristics of QD lasers (Fig. 14). More than sevenfold decrease in the threshold current is observed for a structure based on 10 QD stacks.

TEMPERATURE CHARACTERISTICS

Ultrahigh temperature stability of threshold current was predicted for ideal QD lasers (infinite barriers, no fluctuation of QD sizes, etc.) to be one of the main advantages over conventional (QW lasers.[4] Usually, the temperature dependence of the threshold current density (j_{th}) is described by the characteristic temperature defined as

$$\frac{1}{T_0} = \frac{1}{j_{th}} \frac{\partial j_{th}}{\partial T}$$

In actual laser structures there are several reasons for j_{th} to be temperature dependent:

- The presence of carriers in the OCL and their radiative and nonradiative recombination processes give rise to an additional component of the threshold current, which is governed by the thermal escape of carriers from QDs and depends exponentially on temperature.[9]
- Inhomogeneous broadening of QD array. This effect is similar to that due to recombination in the OCL in the sense that the inhomogeneous line broadening is associated with undesired pumping of nonlasing QDs. So long as the electron and hole populations in the nonlasing QDs are in equilibrium with those in the active QDs, the fraction of threshold current arising from the recombination in nonlasing QDs depends on temperature, and the characteristic temperature is no longer infinite.[25]
- The charge neutrality violation in QDs makes the threshold current component, associated with the recombination in QDs, to be also temperature dependent.[26]
- Small energy separation between the discreet hole levels.[27]
- Temperature dependence of free carrier absorption.

Because of these effects the threshold current and its temperature stability depend dramatically on the parameters characterizing the QD array (surface density of QDs, QD size dispersion, and carrier localization energies in QDs) as well as on the structure design (p-type modulation doping in the active region, thickness of the OCL, doping profiles in the cladding layers and in the OCL, the band offsets at the interface between the OCL and the cladding layers, and the cavity length).

Most of the theoretical works simulate temperature dependence of the threshold current under the assumption of either equilibrium or nonequilibrium carrier distribution in the QD array. Recently, it has been shown that transition from nonequilibrium to equilibrium distribution can result in a very unusual behavior of temperature dependence of the threshold current unique for QD lasers. Fig. 15 shows the temperature dependence of the threshold current density for a high-performance QD laser.[28] At $T<T_1$ ($T_1=150$ K in our case), threshold current density j_{th} only weakly increases with the temperature. This

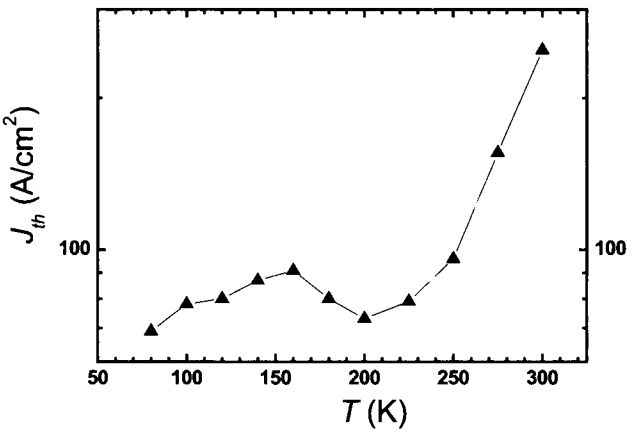


Fig. 15 N-shape temperature dependence of the threshold current density for a laser with narrow QD size dispersion and low threshold current density. Cavity length is 2 mm.

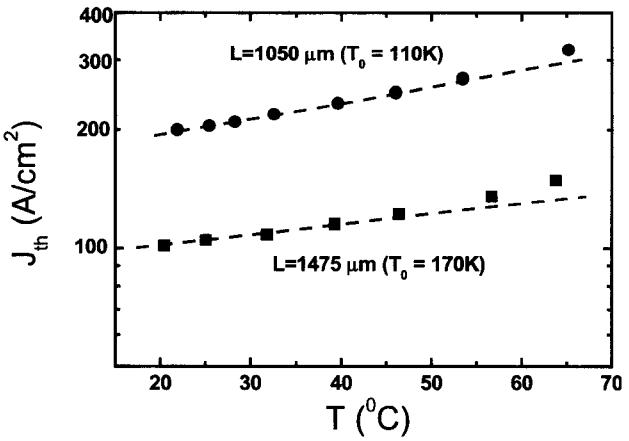


Fig. 16 High-temperature stability of the threshold current for a high-performance QD laser emitting at 1.3 μm. The active area is based on five QD stacks.

increase can be attributed to the presence of carriers in the OCL and their radiative and nonradiative recombination. Then at $T>T_1$, threshold current density decreases with the temperature. Thus at $T_1<T<T_2$ (T_2=200 K in our case), the temperature dependence of the threshold current is described by a negative T_0.[29] Finally, at $T>T_2$, threshold current density starts to increase, with the temperature steeper as compared to low temperature range. Temperatures T_1 and T_2 depend on the confinement of electron and hole energy levels with respect to the continuum and inhomogeneous broadening of QD arrays. For QD arrays with weak confinement and larger inhomogeneous broadening, T_1 can be very low and was not observed in the early experiments.[29]

For commercial applications it is essential to have a high value of T_0 at temperatures above 300 K. Several approaches have been proposed to improve the T_0 values such as placing of QDs in a QW,[30] increasing the bandgap of the matrix,[29] etc. However, even in some cases the T_0 values were increased (up to 350 K in Ref. [30]) threshold current density was also high. First high characteristic temperature operation (160 K) at ambient temperature (below 40°C (313 K)) together with low threshold current density (70 A/cm^2 for 3-fold stacked QDs) was realized in Ref. [31] for a design with negligible external losses in pulsed mode.

Very high characteristic temperature in QD lasers (T_0=230 K up to 80°C (353 K)) was recently demonstrated by applying a QD p-modulation doping technique;[32] the approach, which was earlier proposed to improve the performance of QW lasers.[33] The penalty of p-doping was the reduced differential efficiency (<20% in Ref. [32]). Recently, using undoped 5 fold-stacked QDs, it became possible to increase the T_0 value to 170 K (Fig. 16) between 10°C (283 K) and 65°C (338 K) without paying a penalty of increased internal losses.[34] Remarkably, this device simultaneously showed low threshold current density (100 A/cm^2, 1.5 mm cavity length, uncoated) and high differential efficiency (85%).

TIME-RESPONSE

The modulation bandwidth of QD lasers was measured by different groups and is typically 5–10 GHz.[35–39] The high-speed performance of QD lasers was shown to be limited by the long ''quantum capture'' time of injected carriers.[40] The value of this parameter, which includes transport through the heterostructure, continuum to bound capture time, and intersubband carrier relaxation, varies with injection current and the number of dot levels in the active region. The carrier capture time measured in Refs. [40] and [41] quantitatively agrees with the modulation bandwidth value of 5 GHz. More direct femtosecond differential transmission measurements have been made by using the pump-probe technique.[35] These measurements have been done as a function of temperature (4–300 K) and excitation level (number of carriers per dot). It was shown that, in addition to intersubband electron relaxation in the dots, carrier reemission to the barrier and wetting layers and the density of available states in these regions also play significant roles in determining the carrier dynamics, particularly at elevated temperatures. It is therefore worthwhile to use QD with high confinement energy to suppress carrier reemission from QDs to the WL and OCL. Indeed, in QD amplifiers the use of QDs with high confinement energy (large-sized QDs) resulted in a significant improvement in gain dynamic.[42]

The ''hot-carrier'' problem (or gain compression) can be overcome by injecting electrons directly to the lasing states[43,44] and thus bypassing the usual injection process over the heterojunction barriers. By tunnel injection, ''cold'' electrons are introduced into the lasing subband of the active (gain) region (resonantly or phonon assisted) at or near the Fermi level at a rate higher than the stimulated emission rate, and the electron distribution remains quasi-Fermi even at large drive currents. The hole thermalization rates are usually very large because of the high density of states and band mixing. Hot-carrier effects, mentioned earlier, can therefore be minimized. By utilizing tunnel injection of electrons, enhanced small-signal modulation bandwidth, $f_{-3\ dB}$, and reduced temperature sensitivity of the threshold current, characterized by T_0, were measured in $In_{0.4}Ga_{0.6}As/GaAs$ self-organized quantum dot ridge waveguide lasers. Values of $f_{-3\ dB}$=15 GHz at 283 K and T_0=237 K for 278 K<T<318 K are measured in these devices.[45] However, threshold current density of the laser was high.

Passive mode locking was achieved at 1.3 μm in oxide-confined, two-section, bistable QD lasers with an integrated intracavity QD saturable absorber.[46] Fully mode-locked pulses at a repetition rate of 7.4 GHz with a duration of 17 psec were observed under appropriate bias conditions. No self-pulsation accompanied the mode locking. These results suggest that a carefully designed QD laser is a candidate for ultrashort pulse generation.

LINEWIDTH ENHANCEMENT FACTOR

The linewidth enhancement factor α is a key parameter to characterize the dynamic behavior of semiconductor lasers. It is usually defined as $\alpha = 4\pi/\lambda(dn/dN)(dg/dN)^{-1}$, where N is the carrier density, n is the refractive index, and g is the optical gain. In narrow stripe lasers, large values of α can result in antiguiding, whereas in broad area devices this leads to self-focusing, filamentation, and chirp under modulation. Typically, α is about 2 for InGaAs single QW lasers at carrier densities corresponding to threshold.[47] A record low value of $\alpha = 0.5$ has been reported.[48]

The linewidth enhancement factor can be calculated from the gain spectrum using the Kramers–Kronig relations. In the case of a QD laser with a dot ensemble showing a perfect Gaussian energy distribution and only one energy level for electrons and holes, the gain spectrum is perfectly symmetric around the peak gain energy. In this case, the differential gain is also symmetric about the peak gain energy at which lasing occurs. Thus the differential refractive index change calculated using the Kramers–Kronig relations is exactly zero at the lasing energy, and the QD laser with a quasi ideally distributed (symmetric) ensemble of QDs exhibits a chirp-free operation.[49] Participation of excited states might cause a nonsymmetric gain curve, which increases the linewidth enhancement factor. A narrow QD distribution and well-resolved energy levels of 1.3-μm QD lasers favor substantial reduction in α. The linewidth enhancement factor below the onset of gain saturation has been measured to be as small as 0.1 in QD lasers,[50] which is significantly lower than values reported for QW lasers.

For nearly all applications large spatial coherence of the laser light is needed to focus the beam. Narrow stripe devices with index guiding that gives rise to fundamental transverse mode operation are limited in the output power by the onset of catastrophic optical mirror damage (COMD).[51] The solution of gain-guided broader area devices has been fraught with the problem of filamentation and higher order modes, which destroy the spatial coherence.[52] As beam filamentation strongly depends on the linewidth enhancement factor (α-factor),[53] QD-lasers, which are conjectured to show a reduced α-factor,

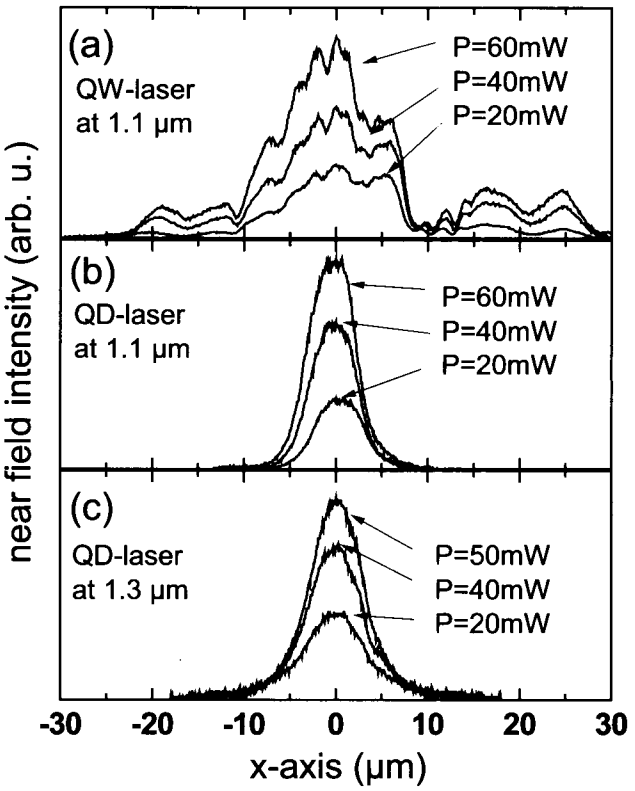


Fig. 17 Near field cross sections in horizontal direction (slow axis) for 6-μm stripe QW laser emitting at 1.1 μm (a), QD laser emitting at 1.1 μm (b), and QD laser emitting at 1.3 μm (c). Cavity length is 1.3 mm in all cases.

are expected to show suppressed filamentation compared to QW lasers.

Recently, near-fields and beam-qualities (M^2) were compared for a narrow stripe 1.1-μm emitting QW laser, a 1.1-μm emitting QD laser, and a 1.3-μm emitting QD laser.[54] The cross section in the lateral direction (slow axis) of the near fields for output powers of 20, 40, and 60 mW is shown in Fig. 17a for the QW laser and in Fig. 17b and c for the QD lasers emitting at 1.1 and 1.3 μm, respectively. The near fields of the QD lasers are Gaussian for all output powers up to 60 mW. In the QW laser, filamentation is present already at an output power of 20 mW. With increasing output power, filamentation increases with more sharp peaks appearing in the beam profile. Additionally, for higher drive currents side lobes next to the laser ridge emerge, too. The beam qualities (M^2) were studied as a function of stripe width for QW and QD lasers.[54] At an output power of 20 mW, M^2 increases for the QW laser from a value of 2.6 for a 3-μm stripe laser to 4.7 for a 10-μm stripe laser. Thus even for a narrow stripe width of 3 μm no lateral fundamental mode emission could be achieved for this gain-guided QW laser. In the case of the QD lasers the situation is quite different. Lateral fundamental mode emission ($M^2 \leq 2$) is demonstrated up to a stripe width of 8 μm for the 1.1-μm QD

laser and up to 9 μm for the 1.3-μm QD laser at an output power of 20 and 45 mW, respectively.

REDUCED CARRIER LATERAL TRANSPORT

Three-dimensional carrier localization in QDs leads to a dramatically reduced lateral carrier spreading in the active area of QD lasers. Carrier capture time from the GaAs matrix to the wetting layer was measured to be less than 1 psec for QDs emitting at long wavelengths.[55] The capture from the wetting layer into the QDs is also very efficient.[56] The fabrication of single-mode lasers may require etching through an active region. For QW devices this would lead to nonradiative recombination of carriers within a diffusion length at the processed surfaces. In contrast, in a QD structure, carrier transport in the lateral direction is suppressed, which makes the surface recombination insignificant. Thus leakage currents in narrow stripe QD lasers are expected to be lower that those in QW devices. In Ref. [56], bright photoluminescence at 300 K was demonstrated from very small 0.2-μm mesas at room temperature, indicating that using QDs as an active medium indeed permits a dramatic reduction of the surface recombination in deeply etched devices.

In QD lasers, lateral transport toward defects in the active area and, correspondingly, nonradiative recombination are also suppressed. This results in reduced sensitivity of the performance of QD lasers to defects in the active area. The influence of high-energy proton irradiation on the device properties of QD lasers was investigated in Ref. [57]. Quantum dot lasers demonstrated enhanced radiation hardness in contrast to QW devices, which indicates that using QDs is promising for lasers in hard radiation environment. First degradation experiments on QD lasers demonstrated the potential of QD lasers for increased operation lifetime.[58,59]

HIGH-POWER OPERATION

High differential efficiency, low threshold current, and suppression of filamentation make QD lasers very promising for high-power applications. Recently, light–current characteristics in QW, quantum wire, and QD lasers have been theoretically treated.[60] Quantum dot lasers were shown to have a major advantage over conventional QW laser from the viewpoint of high-power operation. According to Ref. [60], in properly optimized QD lasers the light–current characteristic is linear with both the internal and external quantum efficiency being close to unity up to very high injection current densities (15 kA/cm^2). Output power in excess of 10 W at an internal quantum efficiency higher than 95% was shown to be attainable in broad area devices.

Present broad-area QD lasers emitting in a spectral range of 0.94 –1.3 μm show CW output powers of 3–6 W.[61–64] Fig. 18 shows the pulsed performance of QD lasers based on 10-fold stacked QDs. The cavity length is 1.45 mm and uncoated facets are used. Differential efficiency as high as 85% and a threshold current density as low as 90 A/cm^2 are realized. The characteristic

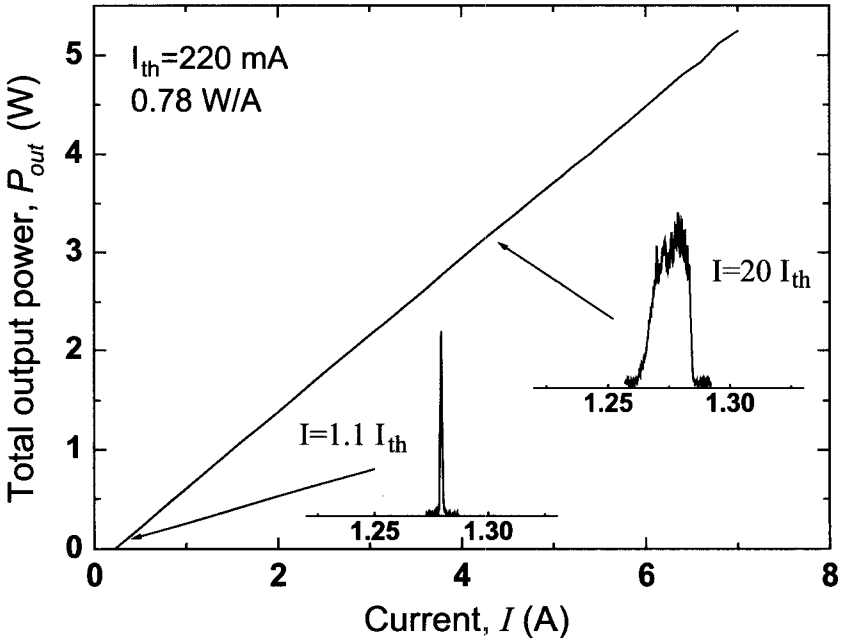


Fig. 18 Pulsed light–current characteristic at room temperature of a broad area QD laser emitting at 1.3 μm. Stripe width is 100 μm, stripe length is 1.5 mm.

temperature is 150 K in a temperature range of 0–70°C. All the key numbers are better than those for commercial InP-based 1.3-μm devices. Narrow 7-μm-wide stripes demonstrated single-transverse-mode kink-free CW operation up to 330 mW for uncoated facets.[65]

VERTICAL CAVITY SURFACE EMITTING LASERS BASED ON QUANTUM DOTS

Vertical cavity surface emitting lasers (VCSELs) emitting at 1.3 μm are very promising for use in high-bit-rate transmission lines for telecommunications. Until recently only the InP/InGaAsP material system has been utilized to fabricate long-wavelength semiconductor lasers because the crystal composition of InGaAsP can be adjusted to produce emitters in the required wavelength region. However, the performance of current long-wavelength lasers still lags behind that of their shorter-wavelength counterparts. The problem limiting the performance of InP-based QW devices is a strong temperature dependence of their characteristics caused by insufficient electron confinement in the active region.[66] This leads to poor temperature stability of the threshold current and emission wavelength, as well as to high values of threshold current densities. Besides, the cost of InP-based devices is substantially higher than that for GaAs-based ones.

An additional disadvantage as far as fabrication of commercial long-wavelength vertical cavity lasers based on traditional InGaAsP–InP heterostructures is concerned is the lack of adequate Bragg mirrors. Mirrors of InGaAsP–InP layers have a small refractive index difference. Thus a Bragg mirror with 50 or more periods is required to achieve the reflectivity value (>0.99) needed for lasing. The introduction of wafer-bonded GaAs/AlGaAs Bragg mirrors to InP-based active regions has led to an improvement in device performance.[67] However, the wafer-fusion technique represents a complex technological process and therefore expensive for large-scale production. In addition, bonded interfaces have low quality, resulting in, for example, increased resistance. Other recent approaches (using top metamorphic AlAs–GaAs DBR, or using AlGaAsSb-based DBRs on InP substrates) do not solve the problem of reliability and cost efficiency, and may be competitive only until the production-oriented GaAs-based technology is developed. Thus for the fabrication of commercial 1.3-μm VCSELs it is essential that the whole structure, including the active area, is grown on the same GaAs substrate, which would enable it to use monolithic high-contrast lattice-matched GaAs–Al(Ga)As or oxidized GaAs–AlO Bragg mirrors.

Two approaches are successfully applied for long-wavelength GaAs-based VCSELs: using InGaAs–GaAs QDs and nitrogen doping of InGaAs layers.[68] Self-organized QDs are particularly advantageous for VCSELs as they offer the possibility to realize very low threshold current density and high temperature stability in properly optimized devices. Owing to suppressed carrier lateral transport very small VCSELs can be fabricated without a deterioration of operating characteristics.

The first 1.3-μm VCSELs grown on GaAs substrate was realized in Ref. [69]. The VCSEL design is schematically shown in Fig. 19. The microcavity is surrounded by (p) and (n) $Al_{0.98}Ga_{0.02}As$ layers (less than $\lambda/4$-thick) followed by 1λ-thick (p) and (n)GaAs current spreading/intracavity contact spacer layers doped to 10^{18} cm^{-3}. Intracavity contacts are used. The spacer layers are followed

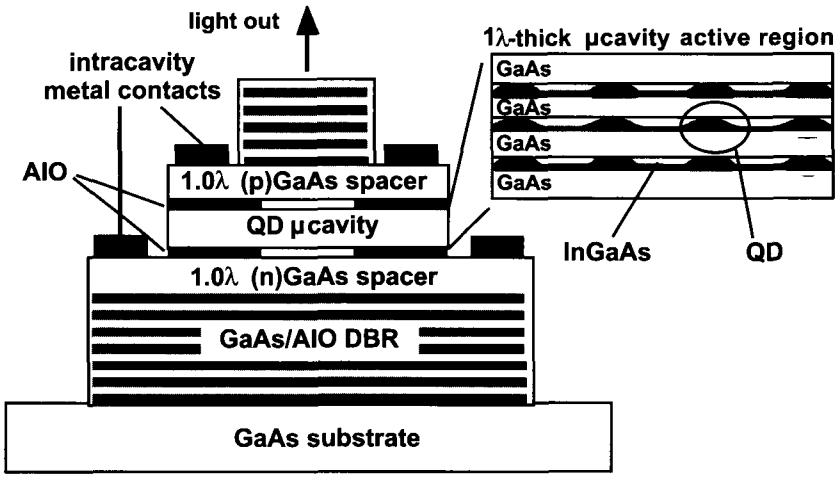


Fig. 19 A schematic cross section of vertical cavity surface emitting laser with QD active region, oxide current aperture, and oxide DBRs.

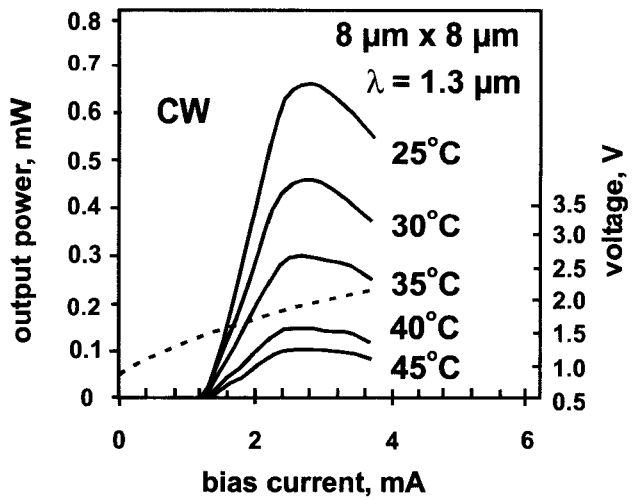


Fig. 20 *L-I-V* curves of a 1.3-μm emitting GaAs-based QD VCSEL at various temperatures.

by DBRs composed of alternating $Al_{0.98}Ga_{0.02}As$ and $\lambda/4$-thick GaAs layers. The $Al_{0.98}Ga_{0.02}As$ layers in the DBR, as well as those surrounding the optical cavity, are selectively oxidized to form Al(Ga)O. The QDs are centered in a 1λ-thick GaAs optical microcavity, whose edges are doped to 10^{17} cm^{-3}. The ends of the microcavity are composed of $Al_xGa_{1-x}As$ linearly graded from $x=0.02$ up to 0.98. The CW light power–current–voltage (*L–I–V*) characteristics of a QD VCSEL are shown in Fig. 20. The threshold current of the QD VCSEL remains practically unchanged with temperature increase. The electroluminescence measurements from a test structure indicate that the lasing proceeds via the QD ground-state transition. The maximum differential efficiency is 64%. The emission wavelength is near 1.3 μm (1.28–1.306 μm depending on the particular position on the wafer). Variation in the threshold current across the wafer is ~10%. The lasing threshold (1.2 mA) is constant down to 0.5-μm oxide apertures, due to reduced carrier lateral diffusion. During the lifetime test in excess of 700 hr CW at 35°C no change in the performance of the device is found.

CONCLUSION

Since the first realization of injection lasing[6] via self-organized QDs, tremendous progress has been achieved in the field of QD lasers. Development of sophisticated growth techniques resulted in arrays of In(Ga,As)–GaAs QD with a relatively small (<10%) inhomogeneous broadening, large confinement energies of electrons and holes, and emission wavelength of about 1.3 μm. Lasers based on such QD arrays demonstrated unique characteristics such as low threshold current density, high temperature stability of threshold current, high differential efficiency, and small linewidth enhancement factor (α-factor). The QDs are also very promising as an active medium for high-performance, cost-efficient, 1.3-μm VCSELs. We believe that further progress in QD growth as well as deeper understanding of the physics of QD lasers will result in the perfection of their characters, reveal new exciting properties, and finally lead to a new generation of high-performance commercial devices for different applications.

ACKNOWLEDGMENTS

This work was done in collaboration with Zh.I. Alferov, D.Bimberg, N.A. Cherkashin, A.Yu. Egorov, I.N. Kajander, P.S. Kop'ev, A.R. Kovsh, J.A. Lott, N.A. Maleev, I.I.Novikov, C. Ribbat, R. Sellin, Yu.M. Shernyakov, and A.F. Tsatsul'nikov. We are very grateful to L.V. Asryan for helpful discussions and kind permission to reproduce parts of his works. We acknowledge the support of the INTAS, the Russian Foundation of Basic Research, the Volkswagen Foundation, and the Alexander von Humboldt Foundation. MVM thanks the Alexander von Humboldt Foundation for the fellowship.

REFERENCES

1. Alferov, Zh.I.; Kazarinov, R.F. *Double Heterostructure Laser, Authors Certificate No 27448, Application No 950840 with a priority from March 30, 1963.*
2. Kroemer, H. A proposed class of heterojunction injection lasers. Proc. IEEE **1963**, *51*, 1782–1784. Submitted October 14, 1963.
3. Dingle, R.Henry, C.H. Quantum Effects in Heterostructure Lasers. U.S. Patent No. 3982207, 21 September 1976.
4. Arakawa, Y.; Sakaki, H. Multidimensional quantum well laser and temperature dependence of its threshold current. Appl. Phys. Lett. **1982**, *40* (11), 939–941.
5. Asada, M.; Miyamoto, Y.; Suematsu, Y. Gain and the threshold of three-dimensional quantum-box lasers. IEEE J. Quantum Electron. **1986**, *22* (9), 1915–1921.
6. Kirstaedter, N.; Ustinov, V.M.; Ruvimov, S.S.; Maxirnov, M.V.; Kop'ev, P.S.; Alferov, Zh.I.; Richter, U.; Werner, P.; Gosele, U.; Heydenreich, J. Low threshold, large to injection laser emission from InGaAs quantum dots. Electron. Lett. **1994**, *30* (9), 1416–1417.

7. http://www.compoundsemiconductor.net/articles/news/6/3/21/1 (accessed March 2003).
8. Bimberg, D.; Grundmann, M.; Ledentsov, N.N. *Quantum Dot Heterostructures*; John Wiley & Sons: Chichester, 1999.
9. Asryan, L.V.; Suris, R.A. Inhomogeneous line broadening and the threshold current density of a semiconductor quantum dot laser. Semicond. Sci. Technol. **1996**, *11* (4), 554–567.
10. Ledentsov, N.N.; Shchukin, V.A.; Grundmann, M.; Kirstaedter, N.; Böhrer, J.; Schmidt, O.; Bimberg, D.; Ustinov, V.M.; Egorov, A.Yu.; Zhukov, A.E.; Kop'ev, P.S.; Zaitsev, S.V.; Gordeev, N.Yu.; Alferov, Zh.I.; Borovkov, A.I.; Kosogov, A.O.; Ruvimov, S.S.; Werner, P.; Gösele, U.; Heydenreich, J. Direct formation of vertically coupled quantum dots in Stranskip–Krastanow growth. Phys. Rev., B **1996**, *54* (12), 8743–8750.
11. Kovsh, A.R.; Maleev, N.A.; Zhukov, A.E.; Mikhrin, S.S.; Vasil'ev, A.R.; Shemyakov, Yu.M.; Maximov, M.V.; Livshits, D.A.; Ustinov, V.M.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D. InAs/InGaAs/GaAs quantum dot lasers of 1.3 m range with high (88%) differential efficiency. Electron. Lett. **2002**, *38* (19), 1104–1106.
12. Maximov, M.V.; Tsatsul'nikov, A.F.; Volovik, B.V.; Sizov, D.S.; Shernyakov, Yu.M.; Kaiander, I.N.; Zhukov, A.E.; Kovsh, A.R.; Mikhrin, S.S.; Ustinov, V.M.; Alferov, Zh.I.; Heitz, R.; Shchukin, V.A.; Ledentsov, N.N.; Bimberg, D.; Musikhin, Yu.G.; Neumann, W. Tuning quantum dot properties by activated phase separation of an InGa(Al)As alloy grown on InAs stressors. Phys. Rev., B **2000**, *62* (24), 16671–16680.
13. Guffarth, F.; Heitz, R.; Schliwa, A.; Stier, O.; Ledentsov, N.N.; Kovsh, A.R.; Ustinov, V.M.; Bimberg, D. Strain engineering of self-organized InAs quantum dots. Phys. Rev., B **2001**, *64* (15), 085305–085312.
14. Liu, G.T.; Stintz, A.; Li, H.; Malloy, K.J.; Lester, L.F. Extremely low room-temperature threshold current density diode lasers using InAs dots in $In_{0.15}Ga_{0.85}As$ quantum well. Electron. Lett. **1999**, *35*, 1163–1165.
15. Maximov, M.V.; Ledentsov, N.N.; Ustinov, V.M.; Alferov, Zh.I.; Bimderg, D. GaAs-based 1.3μm InGaAs quantum dot lasers: A status report. J. Electron. Mater. **2000**, *29* (5), 476–486.
16. Stier, O.; Grundmann, M.; Bimberg, D. Electronic and optical properties of strained quantum dots modeled by 8-band k p theory. Phys. Rev., B **1999**, *59* (8), 5688–5701.
17. Asryan, L.V.; Grundmann, M.; Ledentsov, N.N.; Stier, O.; Suris, R.A.; Bimberg, D. Effect of excited-state transitions on the threshold characteristics of a quantum dot laser. IEEE J. Quantum Electron. **2001**, *37* (3), 418–425.
18. Shoji, H.; Nakata, Y.; Mukai, K.; Sugiyama, Y.; Sugawara, M.; Yokoyama, N.; Ishikawa, H. Lasing characteristics of self-formed quantum-dot lasers with multistacked dot layer. IEEE J. Sel. Top. Quantum Electron. **1997**, *3* (2), 188–195.
19. Maximov, M.V.; Asryan, L.V.; Yu.M.; Tsatsul'nikov, A.F.; Kaiander, I.N.; Nikolaev, V.V.; Kovsh, A.R.; Mikhrin, S.S.; Ustinov, V.M.; Zhukov, A.E.; Alferov, Z.I.; Ledenstov, N.N.; Bimberg, D. Gain and threshold characteristics of long wavelength lasers based on InAs/GaAs quantum dots formed by activated alloy phase separation. IEEE J. Quantum Electron. **2001**, *37* (5), 676–683.
20. Asryan, L.V.; Suris, R.A. Longitudinal spatial hole burning in a quantum-dot laser. IEEE J. Quantum Electron. **2000**, *36* (10), 1151–1160.
21. Park, G.; Shchekin, O.B.; Huffaker, D.L.; Deppe, D.G. Low-threshold oxide-confined 1.3-μm quantum-dot laser. IEEE Photonics Technol. Lett. **2000**, *12* (3), 230–232.
22. Park, G.; Shchekin, O.B.; Csutak. S.; Deppe, D.G. Room-temperature continuous-wave operation of a single-layered 1.3 μm quantum dot laser. Appl. Phys. Lett. **1999**, *75* (11), 3267–3269.
23. Sellin, R.; Ribbat, Ch.; Grundmann, M.; Ledentsov, N.N.; Bimberg, D. Close-to-ideal device characteristics of high-power InGaAs/GaAs quantum dot lasers. Appl. Phys. Lett. **2000**, *78* (9), 1207–1209.
24. Ledentsov, N.N.; Maximov, M.V.; Bimberg, D.; Maka, T.; Sotomayor Torres, C.M.; Kochnev, I.V.; Krestnikov, I.L.; Lantratov, V.M.; Cherkashin, N.A.; Musikhin, Yu.M.; Alferov, Zh.I. 1.3 μm luminescence and gain from defect-free InGaAs–GaAs quantum dots grown by metal-organic chemical vapour deposition. Semicond. Sci. Technol. **2000**, *15* (6), 604–607.
25. Asryan, L.V.; Luryi, S. Tunneling-injection quantum-dot laser: Ultrahigh temperature stability. IEEE J. Quantum Electron. **2001**, *37* (7), 905–910.
26. Asryan, L.V.; Suris, R.A. Charge neutrality violation in quantum dot lasers. IEEE J. Sel. Top. Quantum Electron. **1997**, *3* (1), 148–157.
27. Shchekin, O.B.; Deppe, D.G. The role of p-type doping and the density of states on the modulation response of quantum dot lasers. Appl. Phys. Lett. **2002**, *80* (15), 2758–2760.
28. Novikov, I.I.; Maximov, M.V.; Shernyakov, Yu.M.; Gordeev, N.Yu.; Kovsh, A.R.; Zhukov, A.E.; Mikhrin, S.S.; Maleev, N.A.; Vasil'ev, A.P.; Ustinov, V.M.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D. Temperature characteristics of low

threshold and high efficient quantum-dot lasers emitting in the wavelength range 1.25–1.29 μm. Semiconductor **2003**, *37* (10), 1239–1242.

29. Zhukov, A.E.; Ustinov, V.M.; Egorov, A.Yu.; Kovsh, A.R.; Tsatsul'nikov, A.F.; Ledentsov, N.N.; Zaitsev, S.V.; Gordeev, N.Yu.; Kop'ev, P.S.; Alferov, Zh.I. Negative characteristic temperature of InGaAs quantum dot injection laser. Jpn. J. Appl. Phys. **1997**, *36* (6B), 4216–4218.
30. Maximov, M.V.; Kochnev, I.V.; Shernyakov, Y.M.; Zaitsev, S.V.; Gordeev, N.Yu.; Tsatsul'nikov, A.F.; Sakharov, A.V.; Krestnikov, I.L.; Kop'ev, P.S.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D.; Kosogov, A.O.; Werner, P.; Gösele, U. InGaAs/GaAs quantum dot lasers with ultrahigh characteristic temperature (T_0=385 K) grown by metal organic chemical vapour deposition. Jpn. J. Appl. Phys. **1997**, *36* (6B), 4221–4223.
31. Shernyakov, Yu.M.; Bedarev, D.A.; Kondrat'eva, E.Yu.; Kop'ev, P.S.; Kovsh, A.R.; Maleev, N.A.; Maximov, M.V.; Mikhrin, S.S.; Tsatsul'nikov, A.F.; Ustinov, V.M.; Volovik, B.V.; Zhukov, A.E.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D. 1.3 μm GaAs-based laser using quantum dots obtained by activated spinodal decomposition. Electron. Lett. **1999**, *35* (11), 898–900.
32. Shchekin, O.B.; Ahn, J.; Deppe, D.G. High temperature performance of self-organised quantum dot laser with stacked p-doped active region. Electron. Lett. **2002**, *38* (14), 712–713.
33. Vahala, K.J.; Zah, C.E. Effect of doping on the optical gain and the spontaneous noise enhancement factor in quantum well amplifiers and lasers studied by simple analytical expressions. Appl. Phys. Lett. **1988**, *52* (23), 1945–1947.
34. Ledentsov, N.N. Long-wavelength quantum-dot lasers on GaAs substrates: From media to device concepts. IEEE J. Sel. Top. Quantum Electron. **2002**, *8* (5), 1015–1024.
35. Bhattacharya, P.; Kamath, K.K.; Singh, J.; Klotzkin, D.; Phillips, J.; Jiang, H.T.; Chervela, N.; Norris, T.B.; Sosnowski, T.; Laskar, J.; Murty, M.R. In(Ga)As/GaAs self-organized quantum dot lasers: DC and small-signal modulation properties. IEEE Trans. Electron Devices **1999**, *46* (5), 871–883.
36. Klotzkin, D.; Kamath, K.; Vineberg, K.; Bhattacharya, P.; Murty, R.; Laskar, J. Enhanced modulation bandwidth (20 GHz) of $In_{0.4}Ga_{0.6}As$–GaAs self-organized quantum-dot lasers at cryogenic temperatures: Role of carrier relaxation and differential gain. IEEE Photonics Technol. Lett. **1998**, *10* (7), 932–934.
37. Krebs, R.; Klopf, F.; Rennon, S.; Reithmaier, J.P.; Forchel, A. High frequency characteristics of InAs/GaInAs quantum dot distributed feedback lasers emitting at 1.3 μm. Electron. Lett. **2001**, *37* (20), 1223–1225.
38. Mao, M.-H.; Heinrichsdorff, F.; Krost, A.; Bimberg, D. Study of high frequency response of self-organised stacked quantum dot lasers at room temperature. Electron. Lett. **1997**, *33* (19), 1641–1642.
39. Kuntz, M.; Ledentsov, N.N.; Bimberg, D.; Kovsh, A.R.; Ustinov, V.M.; Zhukov, A.E.; Shernyakov, Yu.M. Spectrotemporal response of 1.3 μm quantum-dot lasers. Appl. Phys. Lett. **2002**, *81* (20), 3846–3848.
40. Klotzkin, D.; Kamath, K.; Bhattacharya, P. Quantum capture times at room temperature in high-speed $In_{0.4}Ga_{0.6}As$–GaAs self-organized quantum-dot lasers. IEEE Photonics Technol. Lett. **1997**, *9* (10), 1301–1303.
41. Heitz, R.; Veit, M.; Ledentsov, N.N.; Hoffmann, A.; Bimberg, D.; Ustinov, V.M.; Kop'ev, P.S.; Alferov, Zh.I. Energy relaxation by multiphonon processes in InAs/ GaAs quantum dots. Phys. Rev., B **1997**, *56* (16), 10435–10445.
42. Borri, P.; Schneider, S.; Langbein, W.; Woggon, U.; Zhukov, A.E.; Ustinov, V.M.; Ledentsov, N.N.; Alferov, Zh.I.; Ouyang, D.; Bimberg, D. Ultrafast carrier dynamics and dephasing in InAs quantum-dot amplifiers emitting near 1.3-μm-wavelength at room temperature. Appl. Phys. Lett. **2001**, *79* (16), 2633–2635.
43. Yoon, H.; Sun, H.C.; Bhattacharya, P.K. Dynamic linewidth of tunneling injection laser. Electron. Lett. **1994**, *30* (20), 1675–1677.
44. Asryan, L.V.; Luryi, S. Tunneling-injection quantum-dot laser: Ultrahigh temperature stability. IEEE J. Quantum Electron. **2001**, *37* (7), 905–910.
45. Bhattacharya, P.; Ghosh, S. Tunnel injection $In_{0.4}Ga_{0.6}As$/ GaAs quantum dot lasers with 15 GHz modulation bandwidth at room temperature. Appl. Phys. Lett. **2002**, *80* (19), 3482–3484.
46. Huang, X.; Stintz, A.; Li, H.; Lester, L.F.; Cheng, J.; Malloy, K.J. Passive mode-locking in 1.3mm two-section InAs quantum dot lasers. Appl. Phys. Lett. **2001**, *78* (19), 2825–2827.
47. Bossert, D.J.; Gallant, D. Improved method for gain/index measurements of semiconductor lasers. Electron. Lett. **1996**, *32* (4), 338–339.
48. Dutta, N.K.; Hobson, W.S.; Vakhshoori, D.; Han, H.; Freeman, P.N.; de Jong, J.F.; Lopata, J. Strain compensated InGaAs–GaAsP–InGaP laser. IEEE Photonics Technol. Lett. **1996**, *8* (7), 852–854.
49. Bimberg, D.; Kirstaedter, N.; Ledentsov, N.N.; Alferov, Zh.I.; Kop'ev, P.S.; Ustinov, V.M.

InGaAs–GaAs quantum-dot lasers. IEEE J. Sel. Top. Quantum Electron. **1997**, *3* (2), 196–205.
50. Newell, T.C.; Bossert, D.J.; Stintz, A.; Fuchs, B.; Malloy, K.J.; Lester, L.F. Gain and linewidth enhancement factor in InAs quantum-dot laser diodes. IEEE Photonics Technol. Lett. **1999**, *11* (12), 1527–1529.
51. Moser, A.; Latta, E.-E.; Webb, D.J. Thermodynamics approach to catastrophic optical mirror damage of AlGaAs single quantum well lasers. Appl. Phys. Lett. **1989**, *52* (12), 1152–1154.
52. Chang-Hasnain, C.J.; Kapon, E.; Bhat, R. Spatial mode structure of broad-area semiconductor quantum well lasers. Appl. Phys. Lett. **1989**, *54* (30), 205–207.
53. Marciante, J.R.; Agrawal, G.P. Nonlinear mechanisms of filamentation in broad-area semiconductor lasers. IEEE J. Quantum Electron. **1996**, *32* (4), 590–596.
54. Ribbat, Ch.; Sellin, R.L.; Kaiander, I.; Hopfer, F.; Ledentsov, N.N.; Bimberg, D.; Kovsh, A.R.; Ustinov, V.M.; Zhukov, A.E.; Maximov, M.V. Complete suppression of filamentation and superior beam quality in quantum-dot lasers. Appl. Phys. Lett. **2003**, *82* (6), 952–954.
55. Graham, L.A.; Huffaker, D.L.; Deppe, D.G. Spontaneous lifetime control in a native-oxide-apertured microcavity. Appl. Phys. Lett. **1999**, *74* (17), 2408–2410.
56. Maximov, M.V.; Sotomayor Torres, C.M.; Volovik, B.V.; Ramushina, E.M.; Skopina, V.I.; Ustinov, V.M.; Gurevich, S.A.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D. Impact of carrier lateral transport and surface recombination on the PL efficiency of mesas with self-organized quantum dots. Phys. Status Solidi, A **2001**, *188* (3), 955–959.
57. Ribbat, C.; Sellin, R.; Grundmann, M.; Bimberg, D.; Sobolev, N.A.; Carmo, M.C. Enhanced radiation hardness of quantum dot lasers to high energy proton irradiation. Electron. Lett. **2001**, *37* (3), 174–175.
58. Lundina, E.Yu.; Shernyakov, Yu.M.; Maximov, M.V.; Kayander, I.N.; Tsatsul'nikov, A.F.; Ledentsov, N.N.; Zhukov, A.E.; Maleev, N.A.; Mikhrin, S.S.; Ustinov, V.M.; Alferov, Zh.I.; Bimberg, D. Long-term stability of long-wavelength (>1.25 μm) quantum-dot lasers fabricated on GaAs substrates. Tech. Phys. **2003**, *48* (1), 131–132.
59. Sellin, R.L.; Ribbat, C.; Bimberg, D.; Rinner, F.; Konstanzer, H.; Kelemen, M.T.; Mikulla, M. High-reliability MOCVD-grown quantum dot laser. Electron. Lett. **2002**, *38* (16), 883–884.
60. Asryan, L.V.; Luryi, S.; Suris, R.A. Internal efficiency of semiconductor lasers with a quantum-confined active region. IEEE J. Quantum Electron. **2003**, *39* (3), 404–418.
61. Zhukov, A.E.; Kovsh, A.R.; Mikhrin, S.S.; Maleev, N.A.; Ustinov, V.M.; Livshits, D.A.; Tarasov, I.S.; Bedarev, D.A.; Maximov, M.V.; Tsatsul'nikov, A.F.; Soshnikov, I.P.; Kop'ev, P.S.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D. 3.9 W CW power from sub-monolayer quantum dot diode laser. Electron. Lett. **1999**, *35* (21), 1845–1847.
62. Kovsh, A.R.; Zhukov, A.E.; Livshits, D.A.; Egorov, A.Y.; Ustinov, V.M.; Maximov, M.V.; Musikhin, Yu.G.; Ledentsov, N.N.; Kop'ev, P.S.; Alferov, Zh.I.; Bimberg, D. 3.5 W CW operation of quantum dot laser. Electron. Lett. **1999**, *35* (14), 1161–1163.
63. Sellin, R.L.; Ribbat, Ch.; Grundmann, M.; Ledentsov, N.N.; Bimberg, D. Close-to-ideal device characteristics of high-power InGaAs–GaAs quantum dot lasers. Appl. Phys. Lett. **2001**, *78* (9), 1207–1209.
64. Klopf, F.; Reithmaier, J.P.; Forchel, A.; Collot, P.; Krakowski, M.; Calligaro, M. High-performance 980 nm quantum dot lasers for high-power applications. Electron. Lett. **2001**, *37* (6), 353–354.
65. Maximov, M.V.; Shernyakov, Yu.M.; Kaiander, I.N.; Bedarev, D.A.; Kondrat'eva, E.Yu.; Kop'ev, P.S.; Kovsh, A.R.; Maleev, N.A.; Mikhrin, S.S.; Tsatsul'nikov, A.F.; Ustinov, V.M.; Volovik, B.V.; Zhukov, A.E.; Alferov, Zh.I.; Ledentsov, N.N.; Bimberg, D. Single transverse mode operation of long wavelength (~1.3 μm) InAsGaAs quantum dot laser. Electron. Lett. **1999**, *35* (23), 2038–2039.
66. Seki, S.; Oohasi, H.; Sugiura, H.; Hirono, T.; Yokoyama, K. Dominant mechanism for limiting the maximum operating temperature of InP-based multiple-quantum-well lasers. J. Appl. Phys. **1996**, *79* (5), 2192–2197.
67. Margalit, N.M.; Babic, D.I.; Streubel, K.; Mirin, R.P.; Mars, D.E.; Bowers, J.E.; Hu, E.L. Laterally oxidized long wavelength cw vertical-cavity lasers. Appl. Phys. Lett. **1996**, *69* (4), 471–472.
68. Choquette, K.D.; Klem, J.F.; Fischer, A.J.; Blum, O.; Allerman, A.A.; Fritz, I.J.; Kurtz, S.R.; Breiland, W.G.; Sieg, R.; Geib, K.M.; Scott, J.W.; Naone, R.L. Room temperature continuous wave InGaAsN quantum well vertical-cavity lasers emitting at 1.3 μm. Electron. Lett. **2000**, *36* (16), 1388–1390.
69. Lott, J.A.; Ledentsov, N.N.; Ustinov, V.M.; Maleev, N.A.; Zhukov, A.E.; Kovsh, A.R.; Maximov, M.V.; Volovik, B.V.; Alferov, Zh.I.; Bimberg, D. InAs–InGaAs quantum dot VCSELs on GaAs substrates emitting at 1.3 μm. Electronics Lett. **2000**, *36* (16), 1384–1385.

Quantum Dots: Electronic Coupling and Structural Ordering

G. S. Solomon
Stanford University, Stanford, California, U.S.A.

INTRODUCTION

Current epitaxial crystal growth techniques, with their precise monolayer (ML) control, have led to abrupt heterointerfaces in III–V and group IV semiconductors. This remarkable heterointerface control is responsible for 1-D carrier confinement in the growth direction: When a thin layer is formed from a more narrow bandgap material in the larger bandgap host, a quantum well (QW) is formed. In the III–V material system, this QW has had a dramatical impact in both semiconductor research and mainstream semiconductor technology. Although significant effort has been concentrated toward extending this control to 2-D and 3-D confinement with quantum wires and quantum dots, the results have been encouraging but not resounding. This is generally because lithography techniques typically used to provide increased lateral confinement do not have the monolayer resolution that is available through epitaxial growth techniques. Furthermore, because of interfacial damage, it is difficult to directly pattern active regions using processing. Although useful structures can be fabricated using surface patterning,[1–4] these techniques are not suitable for all structures. In the 1990s, a purely epitaxial technique was developed to produce quantum dots in the InAs/GaAs and Ge/Si semiconductor systems. This technique utilizes the strain-induced islanding of the Stranski–Krastanow (SK) growth mode, in which the growth surface islands compensate for the increase in energy caused by extra interface surface with a decrease in accumulated strain energy.

The formation of quantum dots by strain-induced islanding has provided a simple, lithography-free method to produce dense ensembles of quantum dots. Unlike the classical self-assembly processes in nonepitaxial systems, in this system, the energies associated with the epitaxial growth process still dominate those that drive the nanostructure formation process. The result is nanostructure features that are not identical but still similar. Strain-induced island formation is perhaps more akin to other surface and interface phenomena such as surface spinodal decomposition,[5] surface reconstruction, and ledge-and-step formation.[6] In fact, our general observation is that the ensemble uniformity and spatial periodicity of InAs islands follow more closely these processes than the self-organized formation of more classical structures such as carbon nanotubes[7] or self-assembled protein structures.[8] Unfortunately, the dominance of the epitaxial process and the large surface migration processes common in this growth lead to large inhomogeneous island size distributions: The spectral features are broadened with respect to QWs, and the narrow, atomic-like transitions are lost in the ensemble broadening. Nevertheless, the association of self-organization with strain-induced islanding, specifically with respect to direct bandgap semiconductors, has done much to focus attention on the possible utility of this system. Although the phenomenological process of strain-induced islanding was observed 60 years ago, the association of spectral features with these islands was made by Tabuchi et al.,[9] whereas the general possibilities of this system have been noted, developed, and championed by Leonard et al.[10] This research is encouraging and lively, and with further fine-tuning of the growth processes, as was needed in the development of the successful QW technology, inhomogeneous broadening can be reduced.

OVERVIEW

The strain-induced quantum dot formation approach has been used effectively to investigate the 0-D structure in nanoscale systems leading to a rich array of results.[11–13] New devices, where the active regions are composed of these quantum dot ensembles, have demonstrated improved characteristics. However, except for the improved thermal stability of lasers demonstrated by Heinrichsdorff et al.,[14] most of these improvements result not from the 0-D confined system but from material compliance. For example, the useful wavelength range of light-emitting diodes (LEDs) and lasers based on GaAs substrates has been extended, but this is not a result of increased carrier confinement. In addition, nearly all of these results rely on the random arrays of islands with slightly varying sizes, and the quest for atomic-like ensembles, either ordered or random, has not been demonstrated to date. This is clearly

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013813

because of the less dominant, secondary role of the energies that drive this ordering process in epitaxial systems.

However, there are benefits to the integration of self-ordering processes with epitaxial deposition. Creating active regions away from processing interfaces and placing these regions within doping and barrier structures while utilizing semiconductor processing and fabrication techniques are attractive approaches to nanostructure fabrication. To this end, we have investigated ways in which vertical growth precision can be incorporated to develop new structures and improve the size and spatial distribution of the strain-induced quantum dot ensemble. In this chapter, we discuss the effect of epitaxially layering InAs islands with small spacer layers of the GaAs host crystal[15,16] as a technique to create either small, well-controlled, vertically coupled structures, or as a mechanism to create an ordered surface array of quantum dots. Because of the strain distribution present around the InAs islands, GaAs deposited on top of an InAs island layer preferentially migrates away from individual InAs dots, resulting in a planarized growth front. If the deposited GaAs is thin enough so that the strain field is still present on the GaAs surface, when subsequent InAs is deposited, there is preferential migration to regions above InAs islands. In this way, InAs island layers, separated by thin GaAs spacer layers, can be vertically aligned in columns.[17]

Two properties can be optimized using these columns. First, the confined energy states within the InAs quantum dots are no longer isolated when two InAs dots are vertically aligned and separated by only a thin GaAs spacer layer. The energy states couple or extend into the adjacent dots.[15] Second, in a subsurface dot layer, the strain distribution from adjacent InAs dots can interact so that a surface strain distribution results, which is a superposition of strain distribution of the subsurface InAs dots.[18] In regions where island density is larger than some equilibrium value, the strain from individual subsurface islands will interfere so that the surface strain will be a superposition of the subsurface distribution. In regions where the island density is smaller than some critical value, the adjacent strain fields from subsurface islands will not interfere as significantly and the resulting surface strain distribution will have regions unaffected by subsurface islands. By depositing multiple island layers, the surface strain distribution can, in theory, be made uniformly periodic. Because the surface strain distribution resulting from the superposition of subsurface island strain affects the surface diffusion, increased spatial ordering of InAs islands on the surface can result. Furthermore, because the size distribution of dots in a layer is affected by variations in the nutrient field surrounding a stable island by creating a more uniform spatial distribution of nucleation sites, the island size distribution can be made more uniform. Thus subsurface layers of InAs dots can be used to create a more structurally and spatially uniform top layer of quantum dots. When optimized to enhance coupling between vertical dots, we call the structures "vertically coupled quantum dot columns." When optimized to enhance surface quantum dot uniformity, we call the structures "subsurface island superlattices."

Before discussing the details of the vertical quantum dot layers, we discuss two salient features of single-layer InAs quantum dot ensembles: the control of ensemble density and the state of in-plane localization. Both of these effects depend on molecular beam epitaxy (MBE) growth conditions. The "Introduction" provides a general introduction to the formation of single quantum dot layers by this technique, in addition to particular discussions of ensemble density and localization. The rest of the chapter is broken into two sections on multiply quantum dot layer growth. The vertical electronic coupling in quantum dot columns is first discussed, whereas ordering by subsurface island superlattices follows.

SINGLE-LAYER InAs QUANTUM DOT ENSEMBLES

Theory of Island Formation

In homoepitaxial crystal growth of cubic materials, the (100) growth surface is the lowest energy facet plane.[19] If the growth temperature and flux rates are appropriate, adatom attachment on the growth surface is to kink and ledge sites, or at island nucleation regions. Thus crystal growth proceeds by the lateral growth of kinks and ledges, or by the expansion of/at monolayer-scale high islands. Ideally, as one monolayer is filled, new monolayer-high nucleation sites are created, and the 2-D (100) growth surface propagates. In contrast, during heterogeneous crystal growth, as more adatoms are deposited onto the growing surface, this growth surface can go through structural changes that can be as dramatical as the complete loss of epitaxial coherency, or as small as changes in surface reconstruction, surface roughness, or abruptness of a heterointerface.[20–23] In the case of InAs on GaAs, the lattice mismatch is 7.2%. The critical thickness for relaxation by dislocation generation is 45 Å, or approximately 15 unstrained InAs monolayers according to the Matthews–Blakeslee model.[24] This critical thickness is based on the mechanical equilibrium of an existing interfacial dislocation. A more accurate theory, based only on the energy density required to generate (not propagate) an interfacial dislocation, indicates the complete absence of a critical thickness.[25] Although this approach cannot be assumed accurate at such large misfit, it does suggest a significantly smaller critical thickness. Because of the similar InAs and GaAs crystal structures, at least one chemisorbed monolayer of InAs can be assumed to be

stable on the GaAs substrate. Below the critical thickness or after a chemisorbed layer, a metastable phase can exist. This is phenomenologically known as the SK growth regime.[26] In the SK growth of InAs on GaAs, there is a thickness region where excess strain is partially accommodated by surface islanding. This growth regime is a transitional growth mode between a compliant planar growth regime that characterizes ideal MBE growth and a plastically relaxed growth, because as islands grow and merge, the surface area can no longer expand to accommodate the increasing strain energy.

Without kinetic effects, island size and shape are predominantly determined by the Gibbs free energy balance:

$$\mu = \mu_0 + \frac{1}{2}\mathbf{S}_{ijkl}\sigma_{ij}(x)\sigma_{kl}(x)V_0 + \gamma K V_0 \quad (1)$$

where V_0 is the molecular volume, γ is the surface tension, and K is the surface curvature (the inverse of the radius of curvature). $\mathbf{S}_{ijkl}$ is the compliance matrix and $\sigma_{kl}(x)$ is the stress field at x. $S_{ijkl}\sigma_{kl}(x)=\epsilon_{ij}$ is the strain, so that $1/2\epsilon_{ij}(x)\sigma_{kl}(x)V_0$ is the accumulated strain energy. $\gamma K V_0$ is the extra surface energy and μ_0 is the total unstrained planar crystal energy. If the epitaxial lattice mismatch were not present, the strain would be absent and the curvature (K) that minimizes the system energy would be zero—corresponding to a flat surface. Thus the lattice mismatch strain leads to a nonflat growth surface with an island size determined by the minimization of Eq. 1. Entropy-of-mixing terms and reduction in the strain energy because of alloy mixing in the near-surface region are not included in Eq. 1. Because of surface diffusion limitations, the surface curvature will take the form of a periodic roughness $a \cos kx$, where a is the roughness amplitude and k is the roughness wavenumber. If τ and n are the local in-plane and surface normals, whereas x and y denote the global crystal in-plane and normal directions, then by geometry, the stresses on the traction free surface are:

$$\sigma_{\tau\tau} = \sigma_{xx}\cos^2\theta + \sigma_{yy}\sin^2\theta + 2\sigma_{xy}\cos\theta\sin\theta,$$

$$\sigma_{\tau n} = \sigma_{nn} = 0, \quad \text{along the surface} \quad (2)$$

To solve for $\sigma_{\tau\tau}$, an Airy solution of the form:

$$\Phi = \frac{\sigma_y^2}{2} + (A + By)e^{-ky}\sin(kx) \quad (3)$$

is used. The stress on the islanding surface is:[27–29]

$$\sigma_{xx} = \sigma(1 - ka(ky - 2)e^{-ky}\sin(kx)) \quad (4)$$

where the y origin is taken as the mean surface. As more materials are deposited, the amplitude a increases. When a approaches the thickness of the heteroepitaxial layer, the troughs of the roughened surface begin to impinge on the heterointerface and stop because the mismatch is no longer present. The trough regions become extended as more materials are transferred away from the interface region to the peaks of the roughened surface, resulting in distinct islands.[20] The island shape is determined by the details of the facet-dependent surface energy γ. In the InAs/GaAs system, the islands are formed from (136) facets.[30]

Based on Eq. 1, as heteroepitaxial material is deposited, the growth surface remains initially flat until the accumulated strain energy exceeds the energy associated with creating extra surface area. The flat region is called the wetting layer in SK growth, and is inherent in this growth mode. Although the concept of critical thickness mentioned previously is based on a continuum model, there has been theoretical[31–33] and experimental evidence[34] showing dislocation and crack nucleation at island edges, suggesting that the islands formed during growth play a more central role in the transition to relaxed heterogeneous crystal growth.

Experimental Control of Quantum Dot Properties

Growth of dislocation-free islands

Except where noted, the InAs dot structures described in this chapter were constructed from epitaxial layers deposited in a Varian Gen II MBE system. The system uses an As_2 flux that is produced by thermally decomposing the As_4 flux commonly used as the arsenic source in MBE. In addition, the As source has an adjustable valve, so that the V/III flux ratio can be precisely controlled. In this way, we can control the surface diffusivity of In. The V/III beam equivalent pressure ratio is kept at 9 for all layers, except where noted. Except where noted, the InAs island region and all subsequent depositions were conducted at 500°C, as measured by the substrate thermocouple, which corresponds to 457°C using a more accurate optical technique.[35] Where growth temperature is varied, it is reported as thermocouple temperature and is approximately 50°C higher than the actual temperature. Each InAs island layer is composed of the equivalent of 3 ML of planar InAs deposited at a rate varied between 0.1 and 0.2 μm/hr. The GaAs growth rate was 0.2 μm/hr. Photoluminescence (PL) spectra were obtained using an Ar^+ ion laser with samples mounted in a circulation He cryostat, except for the LEDs, which are measured at room temperature.

Atomic force microscopy (AFM) images of the islanding surface for different InAs coverages are shown in Fig. 1.[36] This figure shows the transition in the SK growth regime between isolated 3-D island growth and coalesced island growth at 500°C. Fig. 1a shows isolated islanding at 2 ML of surface coverage. Just below this coverage (2 ML) of InAs, the epitaxial surface is planar.

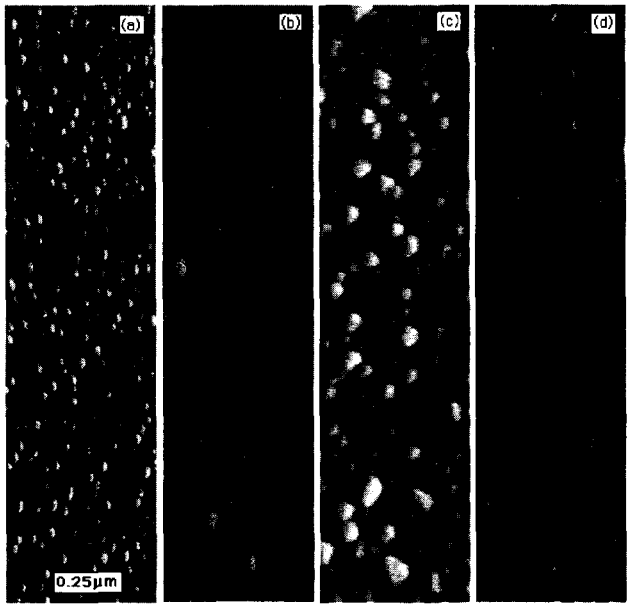


Fig. 1 AFM images of InAs islands formed from various amounts of InAs deposition at 500°C. (a) After 2 ML on InAs, (b) 3 ML of InAs, (c) 4 ML of InAs, and (d) after 6 ML of InAs. The InAs islands are not covered with GaAs, but left bare for AFM imaging. The AFM is conducted ex situ, under ambient conditions.

This abrupt transition to island growth just below 2 ML has been observed by several groups.[9,36,37] With 2 ML of deposited InAs, the average island diameter is 150 Å. Transmission electron microscopy (TEM) investigations[9] indicate that these islands are defect-free. When the equivalent of three planar monolayers of InAs (Fig. 1b) has been deposited, the additional material is accommodated by an increase in island size and density. The island size increases because more materials have been deposited and migrate to the islands. The increase in island density is more complicated and is related to the local wetting layer structure. Because of the lattice mismatch, the wetting layer thickness is only 1.8 ML. Thus very small differences in thickness and strain energy at different sample positions, corresponding to the equivalent thickness on the order of 0.2 ML (0.6 Å), will result in regions that are still growing planarly because this critical wetting layer thickness has not been exceeded, and regions that contain islands because the critical wetting layer thickness has been exceeded. As more materials are deposited, the planar thickness is exceeded over more regions of the sample and the island density increases.

When an additional 1 ML of InAs is deposited, so that the total InAs deposited is 4 ML (Fig. 1c), a significant change in surface topology is evident. The distribution in island size increases significantly and it appears that there are two distinct types of islands. The larger islands are formed by the coalescing of initially isolated islands, which have plastically relaxed through dislocation generation.[34,38] By 6 ML of InAs deposition (Fig. 1d), the transition to a surface topology of plastically relaxed islands is complete.

Fig. 2 shows PL spectra corresponding to the structural transformations observed in Fig. 1. The PL measurements of InAs QWs and quantum dots formed with varying thicknesses of InAs were taken at 8 K using a lock-in technique with a cooled Ge detector for the quantum dot spectra and a Si detector for the QW spectra. The InAs layers have GaAs barriers and are located 500 Å below the sample surface. When 1.5 and 1.75 ML of InAs are deposited, the observed luminescence energy and linewidth are consistent with a uniform InAs (or InGaAs) QW. When the total deposited thickness increases from 1.5 to 1.75 ML, there is a red shift in the spectral peak position because the width of the QW increases. When an additional 0.25 ML is added so that the total deposited InAs is 2 ML, the spectral peak position and shape change dramatically. This is the transition to the quantum dot ensemble structure shown in Fig. 1. This shift occurs when only 0.25 ML (0.75 Å) is added to the surface. The correlation of luminescence structure with the change in physical structure clearly shows that the luminescence originates from the InAs dot regions. In the QD ensembles, the average height is 40 Å and the dots are unalloyed, whereas in the QW, the thickness is less than 40 Å and the well is alloyed with GaAs because of vertical diffusion. It will be shown later in the chapter that significant alloying

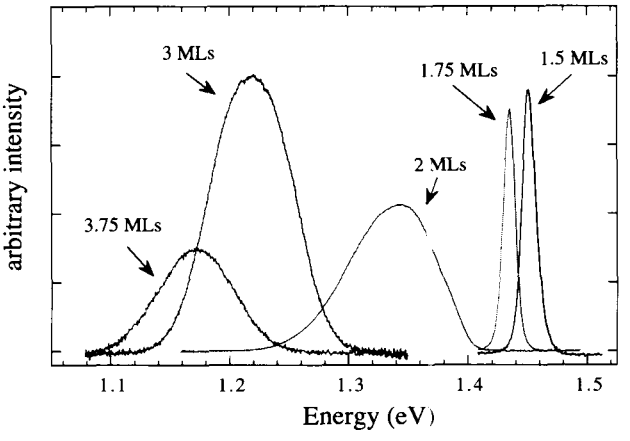


Fig. 2 The 8-K PL of several samples, which contain InAs thicknesses that vary between 1.5 and 3.75 ML. There is a transition from planar growth to dot growth between 1.75 and 2 ML. The intensity decrease between 3 and 3.75 ML is because of nonradiative recombination. (*View this art in color at www.dekker.com.*)

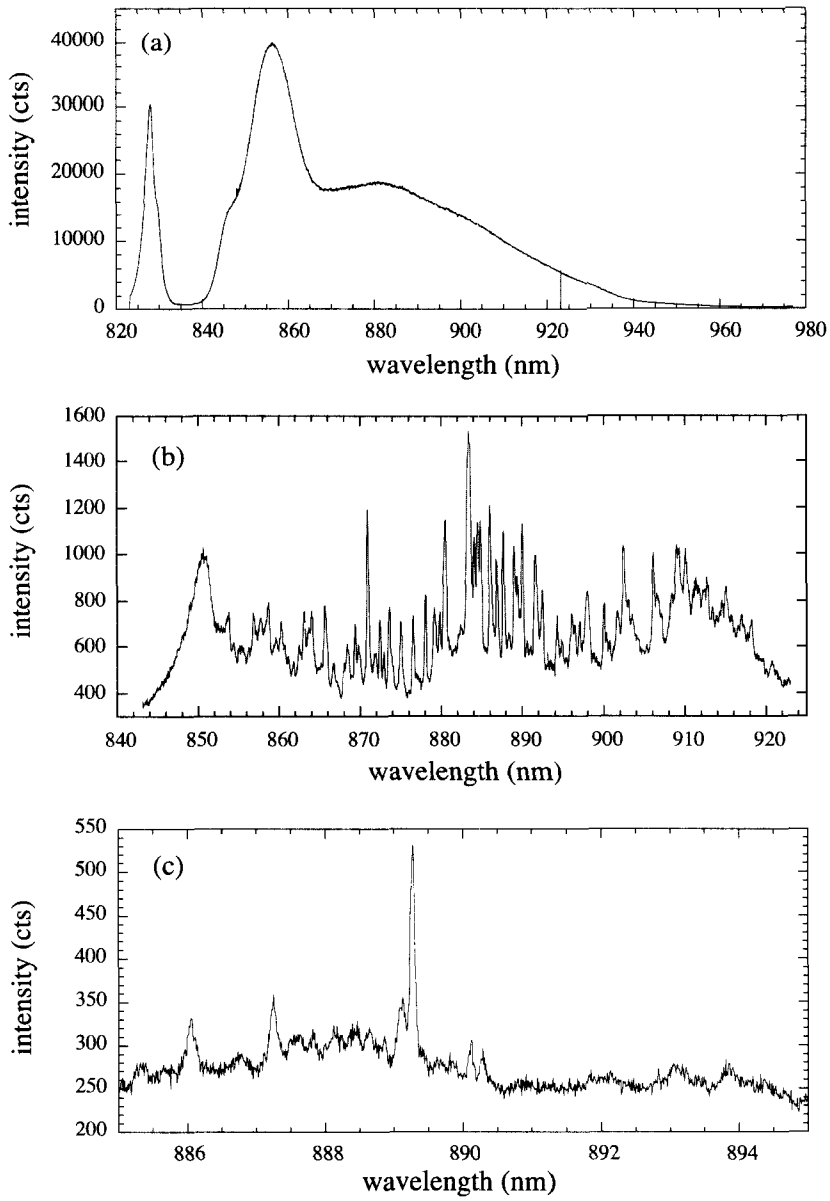


Fig. 3 The 4-K microphotoluminescence of etched posts containing quantum dots. In (a), the unprocessed sample shows the inhomogeneously broadened luminescence from the QD ensemble. The wetting layer is at 850 nm. In (b), luminescence from many individual quantum dots in a 2-μm post is shown. In (c), luminescence in a single quantum dot in a 100-nm post is shown. (*View this art in color at www.dekker.com.*)

is present in the QW region but absent in the quantum dot region, leading to a higher energy bandgap in the QW region than in the QD region. Both the bandgap shift and the change in confinement contribute to the red shift in QD luminescence with respect to the QW luminescence. The QD linewidth increases because of the size distribution in the dot ensemble. The luminescence is attributed to an ensemble of InAs dots of different sizes and shapes that each have very narrow, atomic-like spectral linewidths.[39,40] The integrated intensity also increases when the deposited thickness changes from 2 to 3 ML. This is because of an increase in InAs dot density, as seen in Fig. 1. When the amount of deposited InAs is increased from 3 to 3.75 ML, the spectral peak position again shifts to lower energy because of the increasing size of the InAs islands. However, the integrated intensity now begins to decrease because of defect generation, which acts as nonradiative recombination site.

One of the major issues in the fabrication of quantum dots by strain-induced islanding is that the narrow,

atomic-like emission from single quantum dots is hidden in the inhomogenously broadened ensemble luminescence. Although it is not the purpose of this chapter to investigate the characteristics of single quantum dots, it is instructive to briefly investigate the emission from the isolated quantum dots forming the ensemble. By reducing the number of quantum dots sampled, single quantum dot features can be observed. This can be accomplished by etching posts into the quantum dot sample and observing the luminescence from individual posts. In Fig. 3a, PL from a single InAs quantum dot layer sample is shown. The wetting layer position is 850 nm and the peak of the quantum dot ensemble emission is 885 nm (1.40 eV). Posts isolating the quantum dot layer were fabricated by electron beam lithography and dry etching, after which microphotoluminescence measurements were made on isolated posts. These measurements were made by G. S., M. Pelton, and Y. Yamamoto. In Fig. 3b, microphotoluminescence from a single 2000-nm post is shown. The spectrum shows luminescence from a large number of closely spaced quantum dot features. By reducing the post size, the number of quantum dots within a post is reduced. In Fig. 3c, the post size is reduced to 100 nm and emission from a single quantum dot can be observed. The typical linewidth is 0.9 Å (140 μeV) and is limited by the spectrometer resolution. The linewidth of the wetting layer peak at 850 nm is not significantly reduced because it is a continuous 2-D feature, where the major contribution to luminescence broadening is the thickness variation in the growth direction that fluctuates on spatial scale much smaller than the post diameter.

Substrate temperature

The substrate temperature also affects the InAs QD size and luminescence properties. Fig. 4 shows PL spectra from InAs QD samples where each sample consists of the equivalent of 3 ML of InAs, and only the growth temperature has been varied. The growth temperature was changed from 515°C to 455°C. As the growth temperature is decreased, the spectral peak representing the QD ensemble undergoes a blue shift and the linewidth is reduced. The blue shift in the spectral peak position is because of the reduction in the average dot size. In Fig. 5, an AFM image of the bare dot surface of two samples is shown. The two growth temperatures are 455°C and 500°C. The dot size is reduced from 200 Å to less than 150 Å because of the reduction in In surface diffusion. The dots formed at the lower temperature are less distinct than the dots formed at 500°C. Although the quantum dots are less uniform when grown at 455°C, the spectral linewidth is smaller than in the well-formed QDs grown at 500°C.

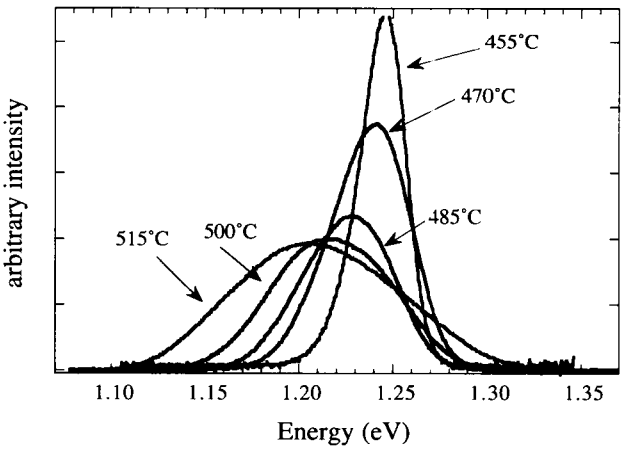


Fig. 4 Spectra of dot structures grown at different growth temperatures. Each sample contains a dot layer composed of the planar equivalent of 3 ML of planar InAs with an undoped GaAs cap. (*View this art in color at www.dekker.com.*)

To explain this result, we first show that the small-scale roughness in the lower-temperature samples does not have a significant effect on luminescence broadening. Assume that a quantum dot exciton has a Hamiltonian of the form $H = H_e + H_h + V_{eh} = H_{ex}$, where H_e and H_h are the Hamiltonians for individual electrons and holes, and V_{eh} is the Coulomb attractive potential. Now assume that in a quantum dot, the individual electrons and holes have a Hamiltonian that is perturbed by roughness potentials $V_{e,d}$ and $V_{h,d}$, respectively. If $H_{e,0}$ and $H_{h,0}$ are the unperturbed QD electron and hole Hamiltonians, the new Hamiltonian is $H = H_{ex,0} + V_{e,d} + V_{h,d}$. Assuming that the exciton is in the ground state, so that $H_0|R_{cm}, 1s\rangle = E_{R_{cm}}|R_{cm}, 1s\rangle$, the variation in the confining potential of the exciton is:

$$\Delta V_{exc,d} = \langle R, 1s|V_{e,d}(z) + V_{h,d}(z)|R, 1s\rangle \tag{5}$$

which we approximate by a combined potential energy component and a geometrical component, so that:

$$\Delta V_{exc} = \Delta V_{rough}\left(\frac{\lambda_c}{a_B}\right)^n \tag{6}$$

where ΔV_{rough} is the change in Coulombic attraction because of local variations in confinement, λ_c is the wavelength of roughness perturbation, a_B is the exciton Bohr radius, and n is a geometrical factor that increases as the confinement goes from 1-D to 3-D. ΔV_{rough} cannot be larger than the Coulomb binding energy, or new confinement regions are formed. In addition, λ_c values larger than a_B lead to inhomogeneous broadening. Thus the maximum broadening occurs when λ_c and ΔV_{rough} are of the order of a_B and the Coulomb binding energy, respectively. This occurs for our case of InAs dots in GaAs

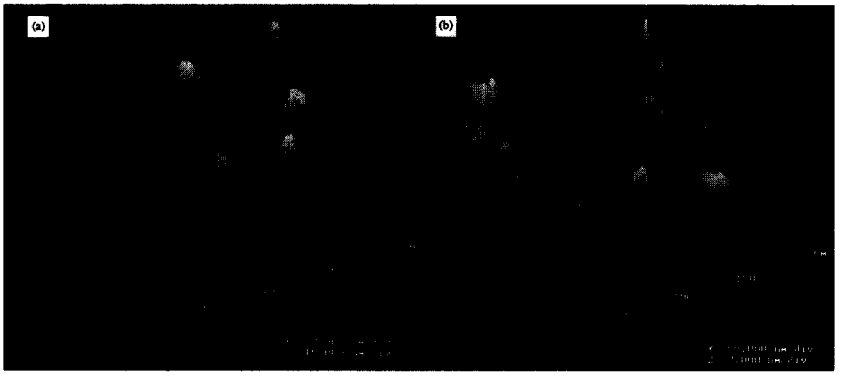


Fig. 5 AFM images of the bare dot surface of two samples identically prepared, except for the growth temperature. In (a), the growth temperature is 500°C, whereas in (b), the growth temperature is 455°C, as measured by the substrate thermocouple. The vertical height scales are indicated and are different in these two images. (*View this art in color at www.dekker.com.*)

when the dot radius equals the exciton radius, which is approximately 100 Å. Because the small-scale roughness that is dominant in the lower-temperature sample is much smaller than 100 Å, it does not significantly contribute to the ensemble luminescence broadening.

In fact, the structure of the lower-temperature dot sample reduces the broadening for a different reason. Because the barriers between the quantum dots are not well defined, the excitons (or electrons and holes) can become partially extended into adjacent quantum dots. This leads to reduced spectral broadening because these carriers interact with an average potential. If $\bar{H}$ is the average Hamiltonian and ΔH_i is the variation from this average seen by carrier i, then the energy variation in ground-state excitons is:

$$\Delta V = \sum_i \langle R_i, 1s|\Delta H_i|R_i, 1s\rangle \tag{7}$$

But this is reduced as the exciton samples more dots (and is eliminated completely in the trivial case when the exciton samples all of the dots). Therefore the broadening decreases as the dots become less defined and the carriers becomes less localized. The reduction in spectral broadening can be further enhanced if acoustic phonon scattering is present between the partially extended states of adjacent quantum dots. This scattering will allow carriers to move to the lowest energy state of the overlapping states. In such a case, the luminescence spectrum will no longer be Gaussian: The higher energy component of the Gaussian distribution will be reduced as carriers relax into the lower energies of the partially overlapping states. Evidence of this can be observed in Fig. 4, where the luminescence spectrum of the sample grown at 455°C is no longer symmetrical.

Although the PL linewidth is reduced and excitons become less localized in any particular dot, this is not coupling from overlapping identical states, nor is this complete delocalization and band formation. The observed narrowing is a result of the weakened in-plane localization, not a result of any increased uniformity in the dot ensemble. Additional confirmation of our model can be seen by investigating the variation in the spectral linewidth with sample (cryostat) temperature during measurement. At low temperatures, the spectral broadening is dominated by inhomogeneous broadening from

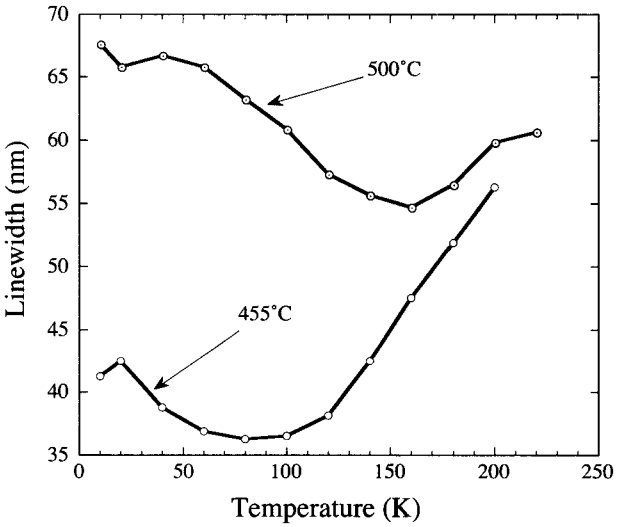


Fig. 6 The spectral linewidth as a function of temperature for two samples fabricated at different growth temperatures. The narrowing in linewidth as the temperature initially increases is because of reduced localization.

the QD ensemble, whereas at high cryostat temperatures, the spectral broadening is dominated by thermal broadening. Fig. 6 shows the results of variable-temperature PL for the two cases described above. An unusual feature in the figure is the initial decrease in spectral linewidth observed in both samples with increasing sample temperature. The linewidth reduction is caused by a decrease in the inhomogeneous broadening as the sample temperature increases. The presence of further spectral linewidth narrowing confirms that complete delocalization is not present. We believe that the increased spectral linewidth reduction is because of variable range hopping between quantum dot regions. This is a thermally activated process involving states of varying energy. We note that thermionic emission above the quantum dot barriers is an unlikely explanation because the barrier between the confined states and the QW wetting region is too large to be of issue at this temperature.

We have confirmed that the QDs formed at low growth temperatures are distinct from the wetting layer QW regions by fabricating QWs with 1.5 and 1.75 ML of InAs. Results of low-temperature PL measurements on these samples are shown in Fig. 7. The data in Figs. 4 and 7 have been combined into Fig. 8. A linear fit is shown between the 1.5- and 1.75-ML QWs grown at 500°C and is extended to the 3-ML region of the figure. Although the 3-ML QD spectral peak position converges to approximately 1.25 eV, the linear fit from the QW region is at a much higher energy; it is clear that the PL peak of the QD ensemble is not converging to the QW (wetting layer) peak position. Although the QD luminescence blue shifts with decreasing growth temperature, the QW luminescence red shifts because

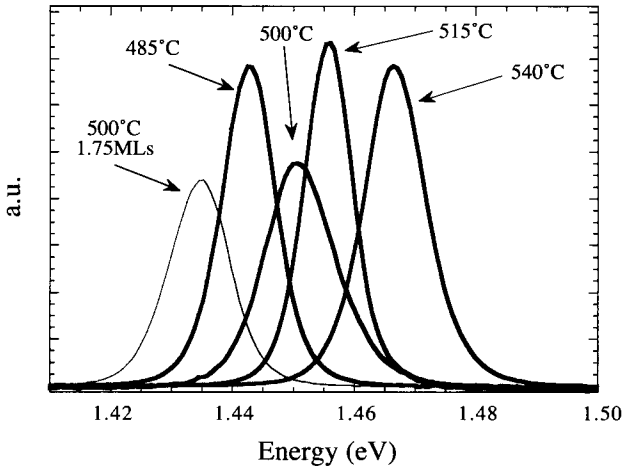


Fig. 7 PL spectra of QWs grown at different growth temperatures with thickness of either 1.5 or 1.75 ML.

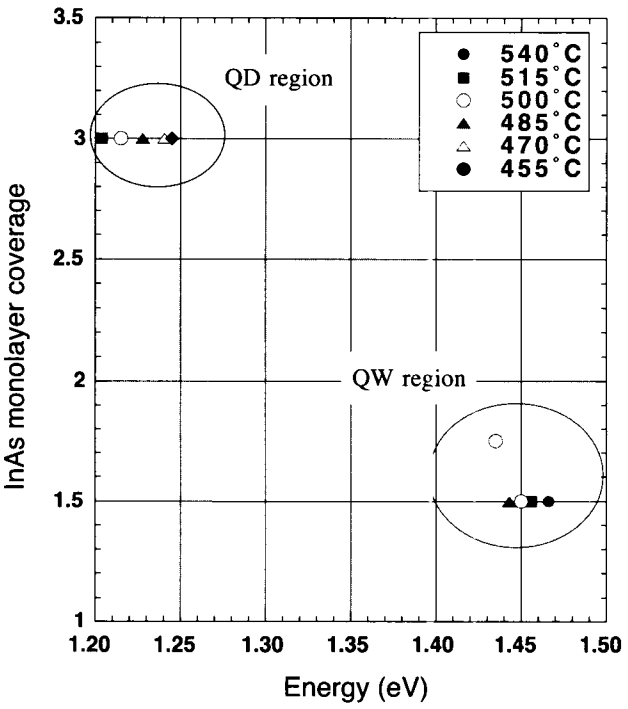


Fig. 8 PL spectra of QWs and quantum dots grown at different growth temperatures indicating that even at low temperature, the quantum dots luminescence is distinctly shifted from the QW luminescence. (*View this art in color at www.dekker.com.*)

of a combination of reduced In desorption and reduced alloying.

V/III flux ratio and quantum dot density

One of the largest determinants of the density of QDs formed using MBE is the V/III flux ratio.[41] Fig. 9 shows AFM images of surfaces where the V/III flux ratio has been varied while the InAs growth rate is maintained constant at 0.19 μm/hr. The V/III ratio is decreased from 36 in Fig. 9a to 18 in Fig. 9b, and finally to 9 in Fig. 9c. As the V/III ratio decreases, the 3-D island density increases, whereas the island diameter remains unchanged. The island density changes from 42% for a V/III ratio of 36, to 49% for a V/III ratio of 18, to 81% for a V/III ratio of 9. In MBE growth, decreasing the V/III ratio increases the In surface migration length. If the island density can be modulated by adjusting the In adatom surface diffusion on the islanding surface, then the island diameter should also change, but this is not observed. However, decreasing the V/III ratio also increases the In adatom surface diffusivity in the wetting layer, allowing it to approach a more equilibrium state with respect to uniform strain and

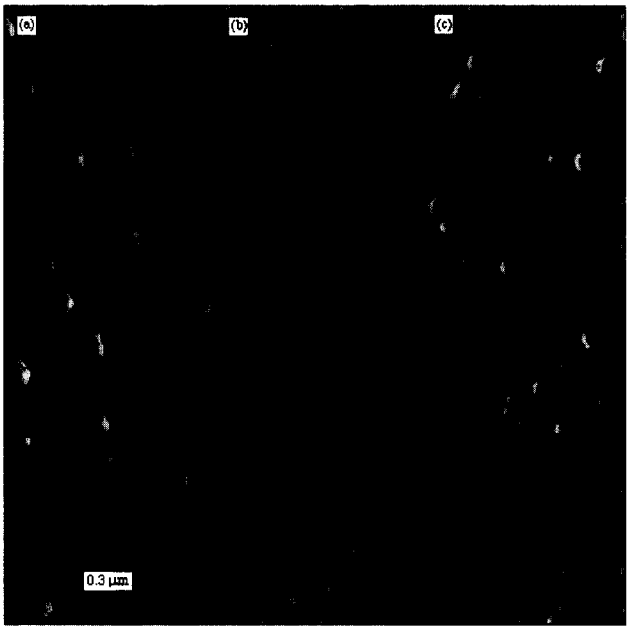


Fig. 9 AFM images of the difference in island density when the V/III flux ratio is changed from 36 in (a), to 18 in (b), and, finally, to 9 in (c). In all cases, the growth temperature was 500°C, the growth rate was 0.19 μm/hr, and 3 ML of InAs were deposited.

uniform roughening. This situation allows for a more uniform and simultaneous transformation to an islanding surface, and hence a more dense island distribution on top of the wetting layer.

Decreasing the V/III flux ratio increases the concentration-dependent surface diffusion in the wetting layer and hence increases the 2-D island size (pill boxes) in the wetting layer. It is important to make clear the distinction between 3-D islands nucleating on top of the wetting layer and 2-D islands contributing to the layer-by-layer growth of the wetting layer. The 2-D islands nucleate between surface steps and coalesce into a layer. In typical MBE growth, their density and hence average size at which they coalesce are affected by surface diffusion: Increasing surface diffusion can lead to larger 2-D islands in the wetting layer. A layer composed of larger 2-D islands will have a more uniform strain distribution than one composed of smaller 2-D islands because of relaxation at island edges; thus this leads to a more uniform transition to an islanded surface and an increased 3-D island density. In addition, the increased surface diffusion will also allow the growth front to roughen more uniformly, and hence more uniformly develop into 3-D islands of increased density. Although two effects within the wetting layer that will increase the 3-D island density on top of this region have been described, it is not known whether one or both of these effects dominate.

Q

ELECTRONIC COUPLING IN VERTICALLY ALIGNED QUANTUM DOTS

Quantum Dot Columns

In addition to the limitations of lithographically defined quantum dots described in the "Introduction," currently, lithography can only be used to define in-plane features and cannot be used to form arbitrary 3-D quantum dot arrays. In this section, we show how arrays of InAs islands in a matrix of GaAs can be vertically stacked and vertically aligned, and we show that such a structure produces electronically coupled quantum dots in the growth direction. Vertical alignment of islands was demonstrated several years ago.[42] These results showed the vertical alignment of two layers in the context of a degenerative roughness process in InGaAs QWs in GaAs, and was believed to be associated with dislocation generation. Here we show the vertical alignment of up to 75 islanding layers with no associated dislocation generation. These vertically aligned islands have been designed to be electronically coupled by using a thin GaAs spacer layer so that the vertical barrier between dots is small. Although these vertical columns do not provide the freedom to make arbitrary 3-D arrays, this technique produces in situ vertical 1-D QD arrays.

In these structures, the thickness of the GaAs spacer layer between InAs islanding layers is defined as the thickness between adjacent wetting layers, and not the thickness between islands. For a 56-Å GaAs spacer layer, the separation between islands in adjacent islanding layers is approximately 15 Å. In Fig. 10a, a high-resolution TEM image of 10 layers of InAs dots shows several vertically aligned InAs dot columns. The sample is tilted slightly off the zone axis to reduce strain contrast, so that the dark regions in the layered structure are predominantly from the increased scattering of the heavier In atoms. The top InAs layer produces much larger, more rounded, and diffuse islands, which we believe are more heavily alloyed with GaAs than the InAs islands in the lower layers. The presence of a new InAs wetting layer close to the peak of the lower islands inhibits mixing of GaAs with the InAs islands. AFM measurements indicate that the islands are approximately 180 Å in diameter, which is confirmed in Fig. 10a. AFM measurements also indicate that the in-plane island size does not change significantly when islanding layers of 10 or fewer are added.[16] Although the small spacer thickness impedes vertical alloying with GaAs, the island shape remains consistent with (113)

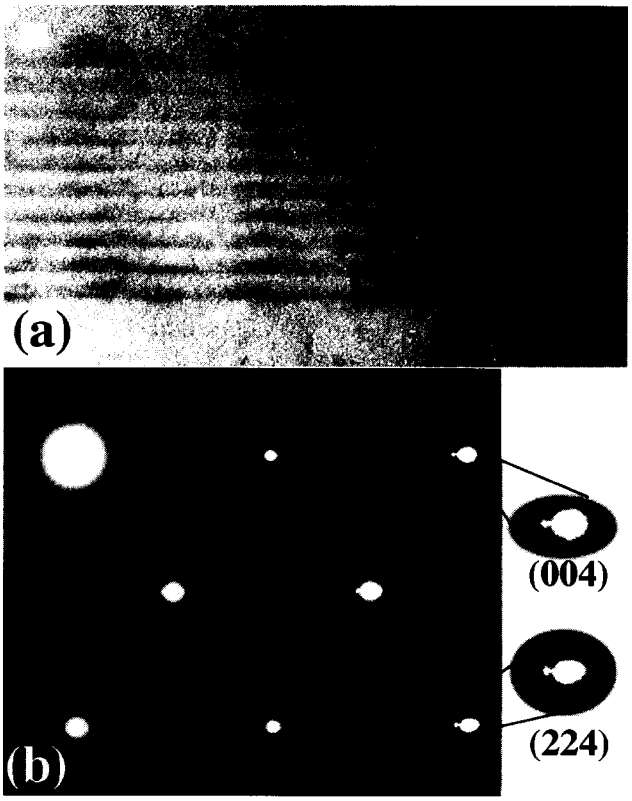


Fig. 10 (a) TEM (110) cross-sectional image of several columns of vertically aligned InAs islands (dark regions). (b) Selected area diffraction pattern of TEM sample in (a) showing superlattice diffraction spots in the (001) growth direction.

sides, as observed by other workers.[43] Several TEM images made perpendicular to the zone axis to enhance the lattice contrast were investigated for dislocations in the layer island regions; none was observed. A search for dislocations at the upper interface between the InAs islands and the GaAs cap was done by tilting the sample up to 30° off the zone axis; again, no dislocations were observed.

A selected area diffraction (SAD) pattern is shown in Fig. 10b. Superlattice and InAs diffraction spots are present at all non-in-plane reciprocal lattice points. Of particular interest are the bright and faint reflections to the left of the GaAs reflections for non-in-plane reflections, such as the (004) reflection. Calculations indicate that the brighter secondary spots are because of the superlattice unit cell and are determined from the average composition and lattice constant of the superlattice. The faint reflections further to the left of the GaAs spot are because of InAs and a higher-order superlattice reflection. In reciprocal space directions that are not the growth direction or the in-plane direction, such as the [224] direction, the InAs reflections correspond to an in-plane InAs lattice constant that is pseudomorphically strained. Because the InAs islands form aligned columns, there are two separate diffracting regions that are structurally connected: the wetting layer regions without columns and the column regions. High-resolution X-ray diffraction (HRXRD) in the [001] direction will be discussed later and confirms that the superlattice diffraction results from the pseudomorphically strained wetting layer region. Because of the high island density, if the islands were relaxed through a dislocation mechanism, this relaxation should extend to the wetting layer region; however, this is not observed.

In Fig. 11, a high-resolution TEM cross section highlighting a single column of InAs islands is shown. It clearly indicates the individual InAs islands and their vertical alignment. The 1.8-ML InAs wetting layer[44] is not observed. X-ray diffraction (XRD) and cross-sectional STM results discussed in "Theory of Quantum Dot Vertical Alignment" show that the wetting layer diffuses vertically; therefore it is very difficult to image with the

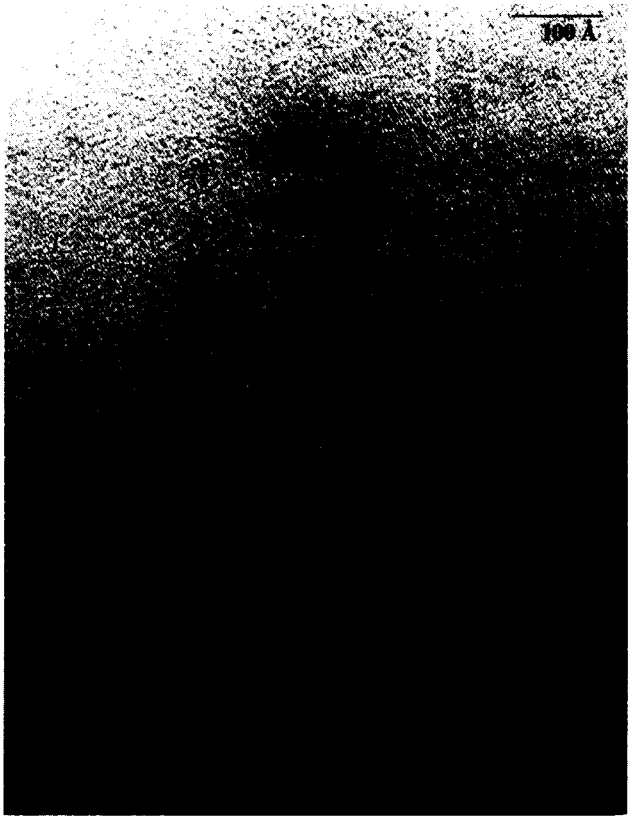


Fig. 11 TEM image showing one column of vertically aligned InAs islands. The island height is approximately 40 Å, and the in-plane dimension is approximately 180 Å.

microscope. The island height is approximately 40 Å, but is difficult to accurately determine because, at the island peaks, only a small number of InAs atoms remain in cross section to contribute to the contrast, and the observed cross section may not intersect the island centers. The lower island interface is flat and abrupt, indicating that the GaAs fills in and smoothes the islanding interface. As in Fig. 10, the island dimensions in the uppermost islanding layer are much larger and, from the contrast change, appear to be alloyed with GaAs. To the right of the island column in Fig. 11 is a four- or five-layer column that does not continue through the upper layers. This decrease in island density after five layers is also observed in our AFM results[16] and is consistent with predictions by Tersoff et al.[18]

Theory of Quantum Dot Vertical Alignment

The vertical aligning of InAs islands is easily understood within the context of an energy balance between interface free energy terms and the lattice mismatch-induced strain energy. The islands form to reduce the large heteroepitaxial free energy component and allow the strain in InAs to partially relax. Because the InAs is partially relaxed, subsequent GaAs deposition is favored in the nonislanding regions, and the valleys between islands will be initially filled. When the islands are covered, the GaAs on top of the islands is locally more stressed than the GaAs in adjacent regions between the islands. When a new InAs islanding layer is initiated, the locally stressed GaAs regions above the old InAs regions become favorable low-energy attachment sites.

The stress field on the islanding surface was described in Eq. 4 as:[27–29]

$$\sigma_{xx} = \sigma(1 - ka(ky - 2)e^{-ky}\sin(kx))$$

The strain is $\varepsilon_i = \mathbf{S}_{ij}\sigma_{ii}$, where $\mathbf{S}_{ij}$ is the compliance matrix. The surface strain is largest on the crest of the islands. Therefore Ga adatoms attaching to the crystal surface at island peaks are more strained than in nonisland regions: There is a strain gradient toward the area between the islands. Thus In adatoms see a potential gradient in the opposite direction. The gradient is expressed as:

$$V = M\sigma_{\tau,\tau} \tag{8}$$

The mobility M is given by:

$$M = \frac{D}{kT} \tag{9}$$

where D is the diffusivity and kT is the thermal energy. The flux J along the surface is proportional to this gradient through the surface adatom concentration c. But if a concentration gradient develops, it also affects the surface flux, so that:

$$J = V\nabla c - D\nabla c = -D\left(\frac{c}{kT}\nabla V + \nabla c\right) \tag{10}$$

If the stress gradient exceeds the concentration gradient, Ga adatoms diffuse to the planar regions, whereas In adatoms diffuse to the island regions.

The local strain relaxation that occurs by islanding in one layer facilitates preferential island formation directly on top of islands in subsequent layers. The islanding process reduces the interface energy by limiting the contact of InAs with the GaAs underlayer, and allows the InAs free surface to partially relax. This partial relaxation is expected to be removed as the InAs island is covered by GaAs. However, if only a thin layer of GaAs is added to the islanding layer, the removal of the partial relaxation will not be complete and the strain field from the buried InAs island will extend to the surface, leaving the GaAs locally strained and possibly distorted. In such a case, the region above a buried InAs island will act as a preferential nucleation site for further islands. Note that the InAs wetting layer and islands form a coupled system, yet SAD and XRD indicate that the wetting layer is unchanged by this preferential islanding and stacking process. The TEM and AFM observations that there is only a marginally small change in the island size with the addition of 10 island layers implies that the modification of the local surface strain state by the buried InAs island is not yet pronounced enough to measurably change the lattice mismatch.

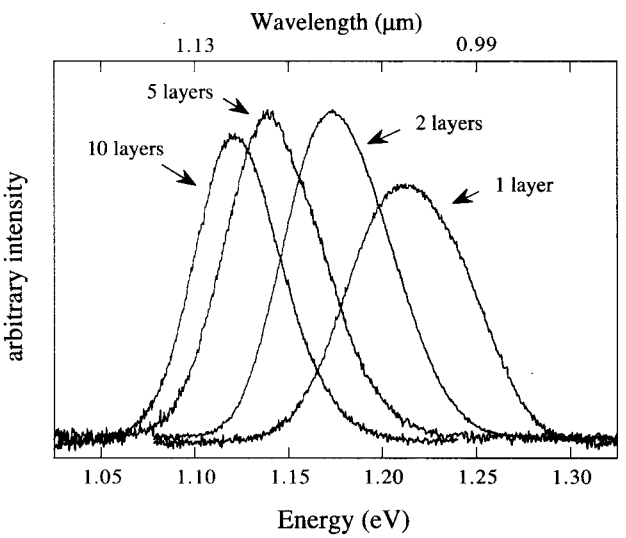


Fig. 12 The 8-K PL of vertically stacked dot layers. The variation in spectral peak position and linewidth is shown when the number of stacked dot layers changes from 1 to 10.

Electronic Coupling in Quantum Dot Columns

Results from 8 K PL measurements are shown in Fig. 12. The single-layer sample is the 3-ML spectrum shown in Fig. 2. Again, the broad luminescence of the single-layer sample is a result of variations in the size of the InAs islands. It is observed that for increasing layers of InAs islands, the spectral peak position shifts to lower energies and the spectral linewidth decreases. In comparison to the single island layer centered at 1.22 eV, the PL peak shifts 73 meV for five layers of islands and an additional 19 meV for 10 layers of islands. Because TEM and AFM measurements show that the change in the island size is small, we attribute the shift in spectral peak position to vertical coupling between InAs islands within a column. There is a 25% reduction in the spectral linewidth between the single islanding layer and the 10 islanding layers. In all cases, the spectral lineshape is Gaussian.

The spectral peak shift and linewidth reduction result from vertical coupling of islands in a column. The tunneling between vertical islands allows carriers to migrate to the lowest energy dot in the column, resulting in a spectral peak shift to lower energy. However, because the number of dots contributing to the spectra is many orders of magnitude larger than the 10 dots in each column, this effect is not responsible for the spectra changes observed here. Vertical coupling also results in spectral peak shifts and linewidth changes by reducing the ground-state confinement energy of each dot column. A miniband is formed; it is the sum of the bonding and antibonding states from the overlap of the individual wavefunctions in a column. The individual dot is approximated as either a sphere or a box with infinite potential barriers so that the linewidth of the ensemble is proportional to:

$$E\sum_i\left(\frac{\Delta\chi_i}{\chi_i}\right) \tag{11}$$

where E is the coupled energy level. $\Delta\chi_i$ and χ_i are the variation of the size and the average size of an island in the ith direction, respectively. The tight-binding formalism is used to model the coupling and resulting energy. In the tight-binding model for a superlattice:

$$E = E_i + S_i + 2T_i\cos(\boldsymbol{q}d) \tag{12}$$

where E_i is the uncoupled energy in each dot, S_i is the energy change in each dot when an adjacent dot perturbs the barrier potential, and $T_i\cos(qd)$ is the wave function overlap term, with dot spacing d and wavevector q. Here the bandwidth $4T_i$ is approximated to increase as $e^{-c(d-1)}$, where c is a fitting parameter weighting the overlap and d is the number of coupled dots. In this analysis, the spectral peak position and spectral linewidth are fitted using the same coupling parameter, and it is assumed that conduction band coupling is dominant. The fit to the peak position and spectral linewidth is shown in Fig. 13. It is found that the peak shift and the reduction in linewidth can be fully attributed to the change in ground-state energy because of coupling.

The strongest coupling state results if all of the dots within a column have the same energy. This coupling will produce the largest spectral peak shift. However, such a state will not result in the largest reduction in the energy linewidth because if all of the dots within a column have the same energy, the in-plane size variations will insure a variation in the lowest miniband state in the ensemble of columns. The weak coupling state, where carriers tunnel through the vertical barriers in a dot column, produces small changes in spectral peak position and linewidth. These changes arise from the increased spatial extent of the dot wavefunctions and not from the strong overlap of coupled states. An interesting case is the intermediate coupling state, in which the vertical variation in the island size in a single column has the same distribution as the in-plane island size variation in an island layer. This coupling state produces the largest reduction in spectral linewidth because all of the dot columns now have the same size distribution. Because it has been shown that a significant element of the in-plane island size variation is caused by small spatial flux nonuniformities across the growth surface,[45] it is unlikely that the vertical distribution in island size can be made equal to the in-plane island distribution. Thus some significant inhomogeneous spectral broadening will always remain, as is the case here.

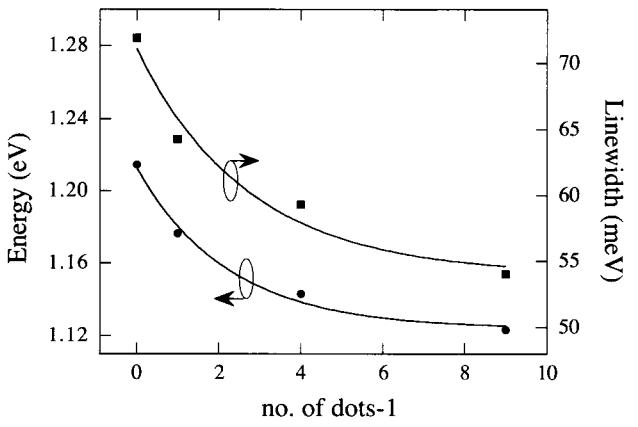


Fig. 13 Changes in (a) spectral linewidth and (b) energy are shown as a function of the number of vertical islanding layers. Both exponential fits use the same fitting parameter.

An additional source of spectral red shift and linewidth reduction is that as the number of island layers increases, the islands may be increasing in size or relaxing. This has not been observed in these 10-layer samples in our TEM or AFM characterization; however, changes in island size are observed in larger columns, as will be seen later in this chapter. If such changes were present here (yet too small to observe) and were responsible for the energy shift, we still do not observe spectral features from individual dot layers in the column, which would result in increased linewidths. Thus even in this scenario, weak coupling (tunneling) from islands in a column to the lowest energy state in the column would still be present.

The coupling in columns can be further investigated by varying the spacer layer thickness between the InAs island layers. The spacer layer thickness in a 10-islanding layer structure was changed between 40 and 50 Å, and 75 and 85 Å. Because these thicknesses are the approximate thicknesses between the wetting layer regions of adjacent islanding layers, the actual vertical spacing between islands varies between 0 and 45 Å. The 8-K PL results for these structures are shown in Fig. 14. The results indicate an increased spectral peak shift in the 40 Å spacer sample as compared with the 50-Å sample. This is consistent with the increased coupling that would be associated with a thinner spacer region. In addition, the spectral linewidth increases when the spacer layer thickness is reduced to 40 Å. The increased spectral linewidth is caused by a disruption in the island structure because the spacer layer

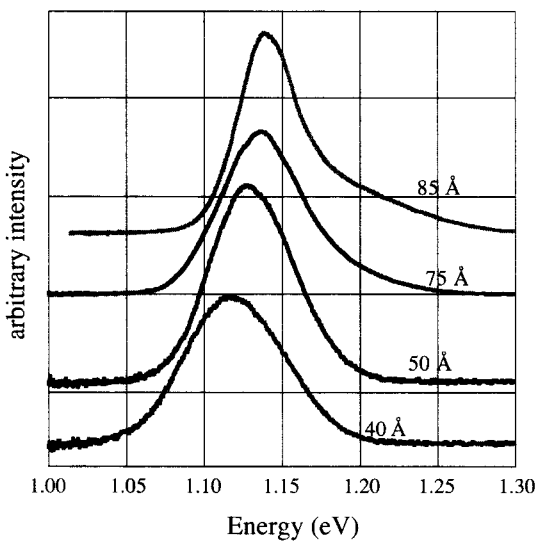


Fig. 14 The effect of varying the spacer layer thickness from 40 to 85 Å is shown. The energy shift is because of the change in electronic coupling. The narrowest linewidth is found in the sample with a 50-Å spacer layer. (*View this art in color at www.dekker.com.*)

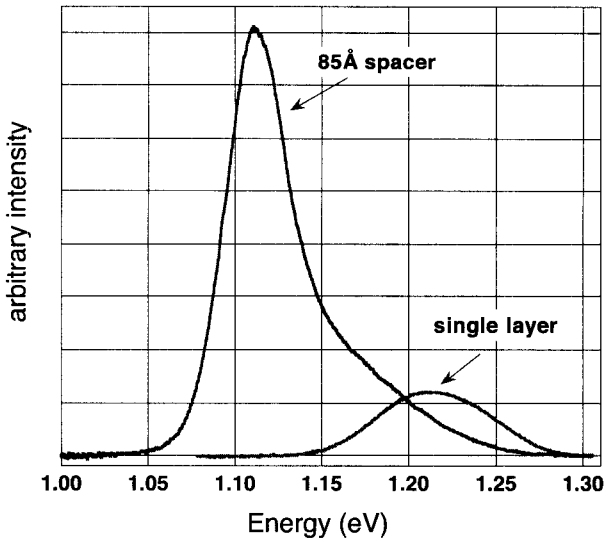


Fig. 15 The 8-K PL of a 10-layer QD sample with an 85-Å spacer is compared to a single-layer sample to show how the larger spacer layer produces a broad high-energy spectral shoulder that indicates incomplete columns. (*View this art in color at www.dekker.com.*)

is now approximately equal to the island height. However, the spectral peak shift is smaller in the sample with the 85-Å GaAs spacer than the sample with the 50-Å GaAs spacer. Here the increased spacer thickness reduces the electronic coupling between islands. In addition, a broad, high-energy shoulder is also apparent in the spectrum with the 85-Å spacer. This shoulder is caused by isolated or only partially aligned islands, where coupling is either not present or is diminished. In Fig. 15, the sample composed of 10 QD layers with 85-Å GaAs spacer regions is shown with the single-layer sample to further illustrate the how the high-energy shoulder corresponds with isolated or partially aligned columns.

We have shown that the wavefunctions of the isolated dots become vertically extended in the vertically aligned dot columns. When the states become vertically extended, the carriers occupying these states become delocalized and sample an average dot potential. This situation is similar to the in-plane delocalization that results when dots formed at low growth temperatures have rough, poorly defined structures (see "Single-Layer InAs Quantum Dot Ensembles" and Fig. 6) and there is in-plane delocalization when the PL temperature is raised, except that the vertical columns are a controlled 1-D chain of dots. We would not expect an increase in the PL measurement temperature to reduce the luminescence linewidth as with the single-layer samples. In Fig. 16, the variation in spectral linewidth with varying sample temperatures is shown for a single InAs QD ensemble sample

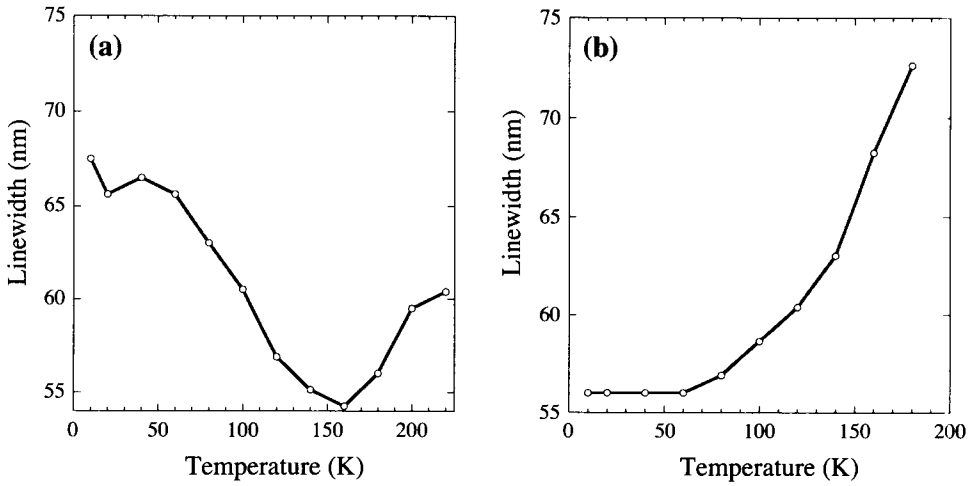


Fig. 16 Variable temperature PL of (a) a single-layer InAs dot sample and (b) 10-layer InAs quantum dots layers. (*View this art in color at www.dekker.com.*)

and one with 10 InAs QD layers. Although the single-layer sample shows the effects of inhomogeneous broadening discussed in "Theory of Quantum Dot Vertical Alignment," the 10-layer sample does not and is predominately thermally broadened.

LEDs Using Coupled Quantum Dots

It is naturally desirable to extend device performance successes gained using QWs in electronic and optoelectronic devices to structures with more reduced density of states (DOS), such as quantum wires and quantum dots. For optoelectronics, the implementation of such structures should result in increasingly efficient LEDs, lower threshold lasers, and novel nonlinear optical devices if: 1) the QD size is reduced to allow for room temperature applications; 2) the QD size uniformity is improved to take advantage of the sharp resonances in the QD DOS; and 3) reliable and simple fabrication techniques are developed. The QD structures described earlier have large-enough confinement energies for room temperature applications; however, the variation in the dot size leads to large inhomogeneous broadening and a loss of the sharp luminescence from the 0-D energy states. We believe that a major advantage of the self-assembled QDs is the high dot densities that can be produced in these self-organized random arrays. As shown in "Electronic Coupling in Quantum Dot Columns," these densities can approach 80% of the ideal close-packed array density, and will be important for optoelectronic device applications.[41] In addition, we believe that the multilayer dot column structure will have an additional impact on optoelectronics. Other groups have made impressive contributions in QD lasers[46] (see the chapter in this book by D. Bimberg), and it is not our desire to describe such results. However, we would like to show in this section how some of the properties of these dot columns can be used in optoelectronic device structures.

The vertical columns of quantum dots described in "Electronic Coupling in Quantum Dot Columns" can be used to increase the optically active volume of LEDs. LEDs are demonstrated in which the active regions are composed of columns of InAs QDs containing either 1, 5, or 10 QDs. Measurements indicate that the single dot layer diode is more efficient than the coupled multidot column diodes. However, at high injection levels, spectra show that much of the luminescence intensity in the single-layer sample is because of the QW formed by the preislanding (wetting) layer and recombination in the doped cladding regions. Luminescence from these regions is nearly eliminated in the multidot column samples. Thus these columnar QD structures may be important to understand high injection conditions, such as lasing. Because of variations in the island size, these diodes are spectrally broad and the sharp spectral features associated with 0-D DOS of these structures are lost. However, of considerable practical importance for optical communications is that spectral peak emission at 1.3 μm has been achieved.[47]

The LED structures consist of an n-type GaAs bottom cladding layer of 0.75μm and a p-type GaAs top cladding layer of 0.29 μm. The diode intrinsic region is 0.11 μm thick and contains an active region of either 1, 5, or 10 InAs island layers in a GaAs matrix. The devices were wet etched to form 20×20 μm mesas. The devices are bottom-emitting and only a portion of the light that is generated in the intrinsic region reaches the bottom interface. The numerical aperture (NA) of the collecting objective is

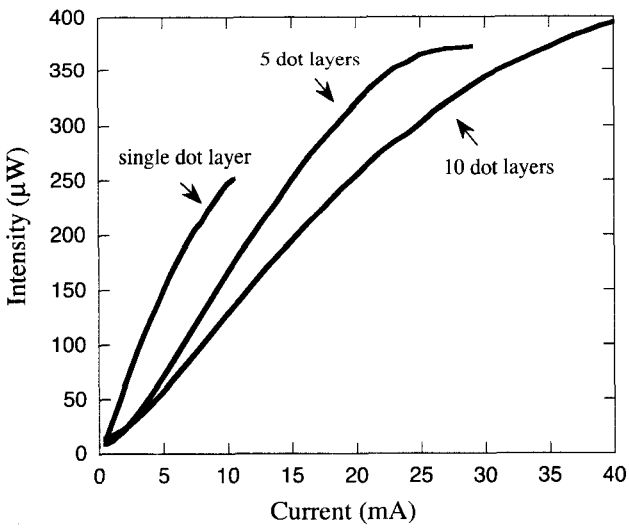


Fig. 17 The light output vs. bias current for three LEDs that use either single, 5, or 10 dot layers at the device active region. (*View this art in color at www.dekker.com.*)

0.25. Therefore neglecting scattered light from different interfaces passing into the objective, only about 1% of the light generated in the intrinsic region is collected by the objective and measured.

Light output collected through the substrate for the three LEDs is shown in the power output vs. current (*L–I*) curves of Fig. 17. Two features are apparent. First, in the linear region of each curve, the slope of the diode with only a single quantum dot layer is higher than the slope of the diodes containing multidot columns. As the number of dots in a column increases, the steady-state carrier concentration throughout the active regions becomes more nonuniform, diminishing the radiative recombination rate. Carrier transport, particularly hole transport, is slowed because the carriers must tunnel through several dots in a column to recombine. Additional evidence of decreased radiative recombination rate is the small flattened onset region in the diodes containing multi-QD columns. If the nonradiative recombination rate is increased, then more current is required to overcome the nonradiative recombination and the slope of the *I–V* curve is initially decreased. The second feature seen is an increase in saturation current and peak intensity. Because there are more dots available, there is now an increased number of states for which recombination can occur. Thus as the number of dots in a column increases, the optically active QD volume increases; this shows the effectiveness of the vertically aligned, columnar QD system. The luminescence in these diodes comes from two distinct regions: the QW wetting layer, and the QDs nucleated on top of the QW. As will be seen, at high currents, a significant portion of the luminescence from the single QD layer diode observed in the *L–I* curve originates outside the QD region, whereas the luminescence is well confined to the QD regions in the diodes containing columns of multiple QDs.

The spectra of diodes containing columns of 1, 5, or 10 InAs QDs are shown in Fig. 18. As the number of QD

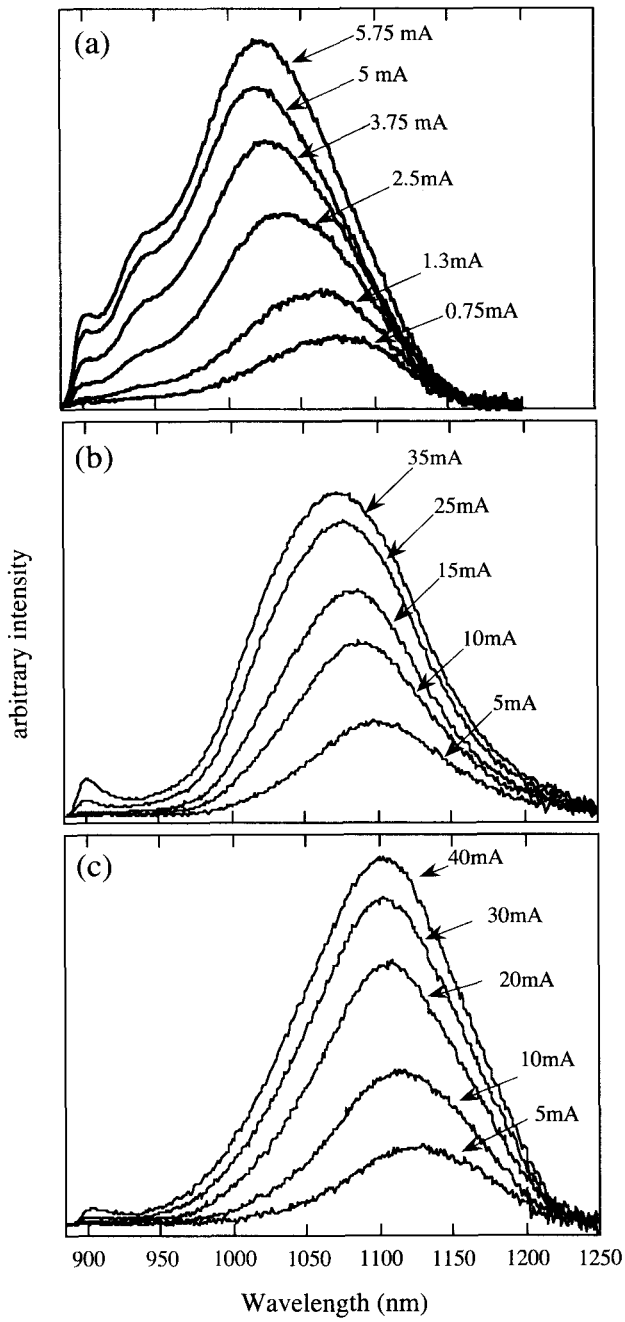


Fig. 18 Luminescence spectra for three diodes with different numbers of quantum dot layers in the active region with various current bias. (a) Single quantum dot layer, (b) columns of five quantum dots, and (c) columns of 10 quantum dots.

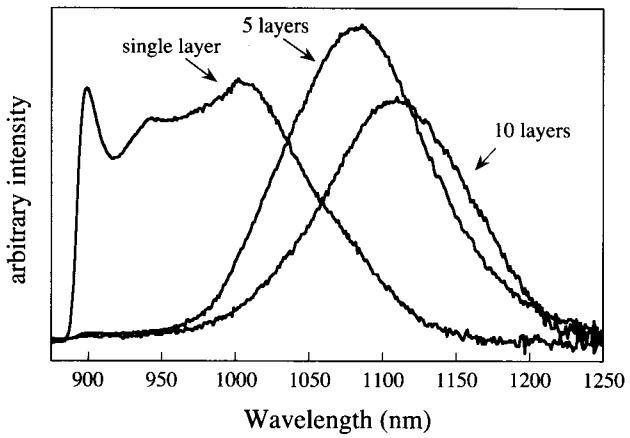


Fig. 19 Comparison of the three diodes with columns of 1, 5, and 10 quantum dots using the same current of 15 mA.

layers increases, the spectral peak position of the diode shifts to longer wavelengths and the spectral linewidth decreases (Fig. 12). As the current increases in each of the diodes, luminescence from the doped GaAs cladding region increases at 900 nm. This luminescence feature is the luminescence shoulder from recombination in the doped region of the diodes where bandgap narrowing is present. From top-emitting LED samples, it is estimated that one-half of the feature (the region below 900 nm) is absorbed by the GaAs substrate. Carriers travel through the active region and recombine with majority carriers in the doped region. The prominence of this feature is significantly reduced as the number of QDs in the dot columns is increased because carriers are more effectively captured in the InAs dot region. The spectra of the diodes with a single QD layer (Fig. 18a) exhibit an additional feature between the main QD peak and the luminescence from recombination in the barriers, which is absent in the diodes containing columns of multiple QDs. This extra peak is because of radiative recombination in the QW wetting layer region. The peak position is well modeled, assuming a strained, coherent InAs wetting layer (QW) with a wetting layer thickness of 1.8 ML, a thickness observed in Fig. 1 and by other researchers.[43,44] The coupling of dots in the QD columns is larger than the coupling between QWs, so that the dominant carrier transport is through the QD columns.

As the current increases in the diodes, the spectral peaks in all samples shift to shorter wavelengths. The LEDs are biased under d.c. conditions. Sample heating is not responsible for this effect because it will reduce both the InAs and GaAs bandgaps similarly and produce a red shift in each. The spectral shifts with bias are larger in the single-layer QD diode, whereas they are equal in the two multi-QD column samples. In all cases, the spectral peak shift stabilizes at higher currents. The spectral peak shifts result when the lower-energy states from larger QDs become saturated. The extra shift in the single QD layer sample results from the proportionally larger luminescence from the QW wetting layer.

A comparison of the three different diodes at the same current of 15 mA is shown in Fig. 19. In the single-layer QD diode, the intensity increase between 6 mA (Fig. 18a) and 15 mA results almost entirely from recombination in the QW InAs wetting layer and the GaAs barrier region; the energy states contributing to the main 0-D spectral peak luminescence are saturating. Thus this luminescence is a mixture of light from the 2-D QWs and light from the 0-D QDs. The luminescence intensity of the diodes containing the multi-QD columns results entirely from the 0-D energy states. At 15 mA, the peak intensity of the diode with five dot columns is larger than the peak intensity with 10 dot columns. Even without scaling, it is apparent that the internal quantum efficiency of the device is reduced between the 5- and 10-dot column samples.

An approximate comparison of internal quantum efficiency can be made by comparing the diodes when the current per QD layer is equal, and the luminescence intensity is scaled by the number of QDs in a column. This is shown in Fig. 20, where the current per dot layer is 3 mA. The normalized peak intensity of the diode with the single dot layer is the highest of all the diodes, so that even with respect to only the 0-D luminescence, the single-layer diode has the highest efficiency. As observed in multiple quantum well (MQW) LEDs, holes are not being evenly distributed throughout the multi-QD column active

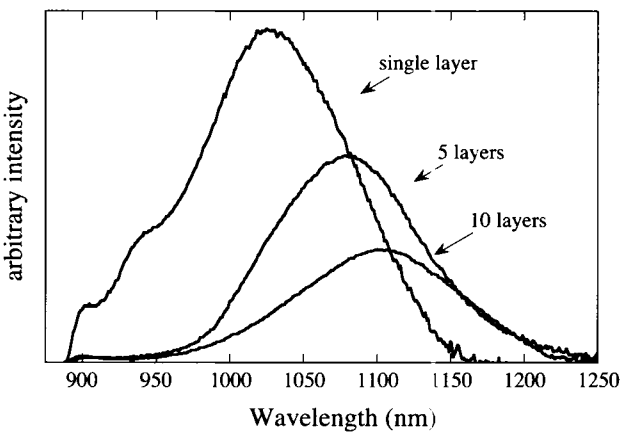


Fig. 20 Comparison of the three diodes with columns of 1, 5, and 10 quantum dots where the current per quantum dot layer is constant at 3 mA for each of the three diodes. The luminescence from each diode is divided by the number of QD layers in each diode.

region because their large effective mass reduces their tunneling probability between QDs in a column. However, the effects of nonradiative recombination processes, such as recombination at crystalline imperfections, must also be considered. This does not appear to be the case because the slope of the *I–V* curves of diodes with either 5 or 10 dot columns is similar, implying that there is no significant increase in nonradiative recombination between the two diodes.

In summary, the increased active volume of the columnar structure containing multiple QDs leads to increased peak luminescence intensity and reduced carrier recombination in the doped GaAs region. However, this comes with a penalty, which is a decrease in luminescence efficiency as the QD column size increases. It is believed that the reduced luminescence is because of a partitioning of holes in the QD columns, but other effects such as nonradiative recombination at structural defects cannot be ignored. Finally, in the diodes containing QD columns, the luminescence from the QW wetting layers regions is eliminated because the coupling within the dot columns facilitates carrier transport through the QD columns. This is a significant advantage to using the quantum dot columns because early laser structures appear to lase through the wetting layer states or excited states of the dots. Because of the hole partitioning in columns with many dot layers, an optimum quantum dot active region should contain columns with a limited number quantum dots. The number will vary depending on the structure, but the results shown here indicate that columns containing about five QDs adequately balance the reduction in QW luminescence without significant hole transport penalties.

IN-PLANE InAs ISLAND ORDERING USING A SUBSURFACE ISLAND SUPERLATTICE

Columns Composed of Many Quantum Dot Layers

The focus of "LEDs Using Coupled Quantum Dots" was the electrical coupling that can occur in vertical columns of a small number of quantum dot layers. A different effect can be observed when a large number of layers are stacked. In this regime, the column structure can increase the structural and spatial uniformity of the uppermost in-plane QD ensemble. Although we have explored some of the structural properties of these columns in "LEDs Using Coupled Quantum Dots" using TEM, we will more closely examine these properties using AFM, XRD, and cross-sectional scanning tunneling microscopy (STM).

A cross-sectional STM image of columns of five InAs QDs is shown in Fig. 21. Here the image size is

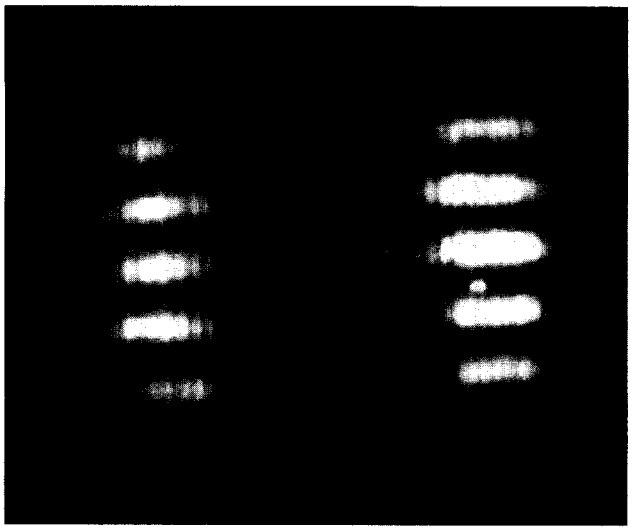


Fig. 21 Cross-sectional STM empty states image of two columns of InAs quantum dots (bright) in a GaAs matrix (dark). The field of view is 600 Å in (110) (horizontal) and in (001) (vertical). (Courtesy of W. Wu and J. R. Tucker.)

600×600 Å. This and other STM images were made by W. Wu and J. R. Tucker at the University of Illinois, Urbana-Champaign. These samples are prepared by in situ cleaving in a 5.5×10^{-11} Torr chamber to expose the (110) surface. The measurements are made at room temperature. The STM image is an empty states image derived from current variations across the surface when electrons tunnel from the STM tip to available states in the valence band. Because of the charge transfer on the (110) III–V unreconstructed surface, empty valence band states are localized on the group III sublattice. Thus this imaging highlights Ga or In atoms. Because of the narrower bandgap of InAs and the larger size of the In atom (which pushes the exposed In further out into the vacuum), the tunneling current is higher in the In-rich regions than in the Ga-rich regions. Hence, the contrast is directly related to the mole fraction of In. In these images, discrete bright areas are the result of In surface atoms. Fainter bright spots are assumed to be from In atoms in the next group III layer below the surface.[48,49]

This image, as well as other STM images, are consistent with the TEM micrographs. The in-plane dot diameters in Fig. 21 are approximately 130 Å and the heights are 30 Å. The spacing between the two columns is 200 Å. As in cross-sectional TEM, the exact diameter cannot be measured directly from these images because the cleave does not necessarily bisect the center of the QDs. TEM and STM are complimentary measurements. The advantage of this STM technique will be seen later in this section: Although TEM images are formed by

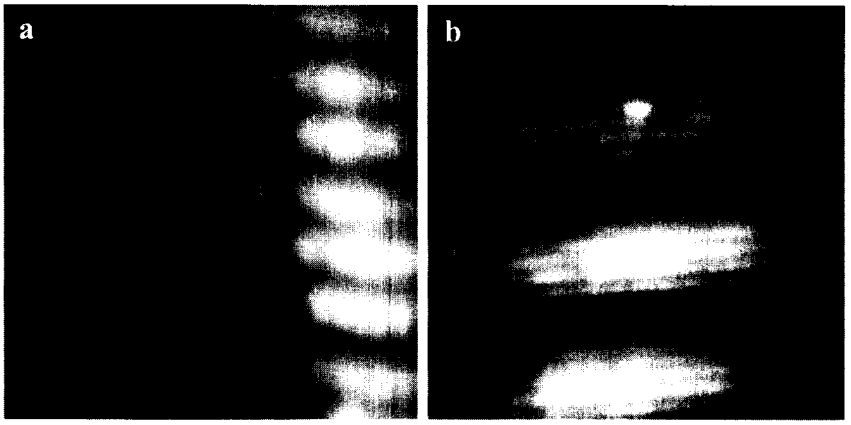


Fig. 22 Cross-sectional STM image of (a) the region between two 20-layer quantum dot columns showing In diffusion between dot columns, and (b) the region around a column indicates no significant In diffusion from the dot regions. (Courtesy of W. Wu and J. R. Tucker.)

columns of atoms in the plane of the image, the STM images are formed from states close to the surface. Thus individual In atoms that have diffused into GaAs regions can be imaged with STM. In addition, the DOS can also be examined using STM by sweeping the tip voltage.

We now focus on uniformity improvements that can be made using large columns containing up to 75 layers of dots. In these structures, there is an interaction between neighboring columns, which imposes a strain field on the growing surface and can lead to dot ensembles of both increased size uniformity and increased spatial uniformity. We begin by showing that, despite the layering of dots to form columns, the InAs diffusion from these dots is limited, especially with respect to the large InAs diffusion observed in the wetting layer regions.

In Fig. 22, a cross-sectional STM image of two adjacent 20-layer columns of InAs islands is shown. As in Fig. 21, the STM image is an empty states image highlighting tunneling current differences associated with Ga or In. In Fig. 22a, a speckled band, which is the planar wetting layer region, is present between the dot columns. This region has been shown by various groups to consist of between 1.6 and 1.8 ML of InAs, determined through reflection high-energy electron diffraction (RHEED). Although it has been shown in earlier InAs planar epitaxy studies that the surface segregation in this layer results in a broadly alloyed region,[48–50] in strain-induced quantum dot research, it is often assumed that this region remains relatively intact. In Fig. 22a, the In diffusion can be easily seen to extend to the height of the InAs dots, which is approximately six unit cells.

We have observed the global structure of this wetting layer region using XRD. In Fig. 23a, XRD results are shown for a sample consisting of 10 InAs dot layers. The XRD was conducted using a Philips 4 crystal X-ray diffractometer. In the figure, we observe superlattice diffraction from a supercell consisting of the GaAs spacer and the InAs. Peaks for the $n=0$, ± 1, and ± 2 reflections are present, with the (004) GaAs substrate peak residing between the $n=0$ and $n=1$ reflections. A fit using a dynamic diffraction model gives (Fig. 23a) a strained InAs thickness of 1.9 ML (± 0.2 ML) and a GaAs thickness of 52 Å. Our interpretation of these results is that the wetting layers (in the dots wetting layer system) are the only contribution to the superlattice diffraction. Because Fig. 23a represents symmetrical diffraction about the (004) GaAs reflection, the dot regions, with their varying InAs and GaAs vertical sections, should not contribute to superlattice diffraction. The only extensive regions with periodic structure in the growth direction are the regions of the sample without InAs dots where only the wetting layers are present. To verify this result, XRD was conducted on a similar sample with only InAs wetting layers. This sample is composed of 10 periods of 1.5 ML of InAs; the GaAs spacer layer is 50 Å. The results are shown in Fig. 23b. When fitted using the same diffraction model, the best fit occurs for 1.5 ML (approximately 9 Å) of strained InAs.

This measured value of 1.9 ML for the 10-layer dot sample is close to the 1.8-ML wetting layer thickness observed by other researchers using RHEED. The RHEED measurements determined the equivalent amount of materials deposited leading to the dot transition, whereas this XRD measurement determines the wetting layer thickness after the dot transition. Because the transition to island growth results in an overall change in the surface topology, it is interesting to note that the wetting layer thickness is unchanged by the transition to islanding.

The width and decay of the superlattice diffraction in Fig. 23 indicate the uniformity of the wetting layers over

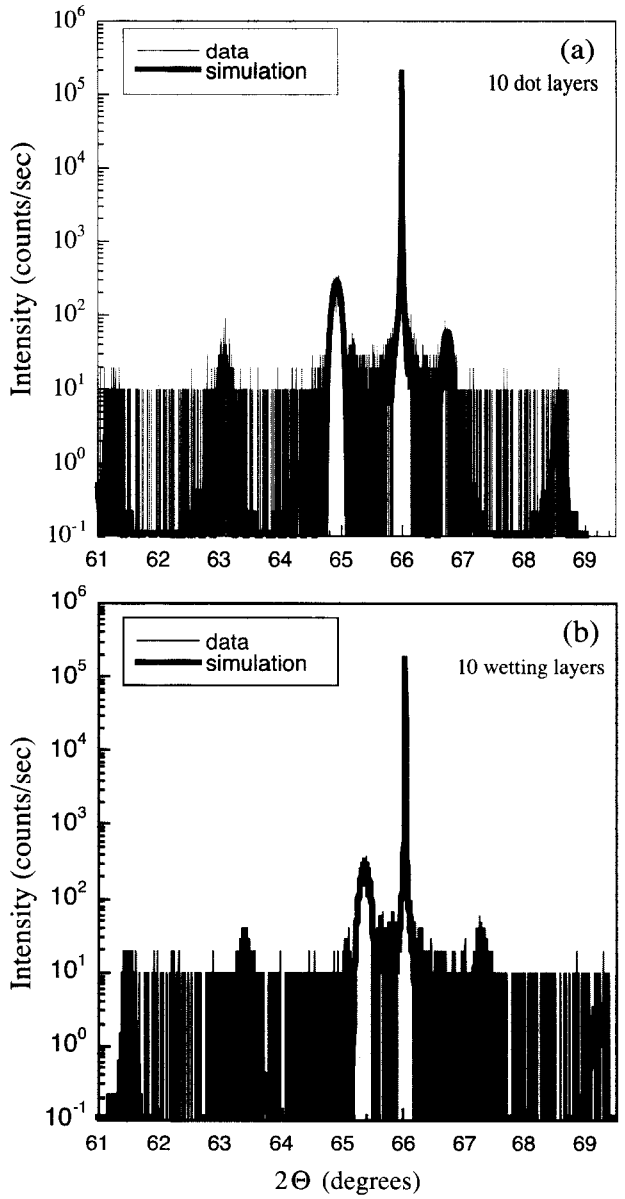


Fig. 23 X-ray diffraction from (a) a 10-dot layer sample formed in the columnar structure, and (b) a QW 10-wetting-layer sample (nominally planar InAs layers). Simulations are shown in black. In both cases, the superlattice diffraction is caused by the wetting layer regions.

the superlattice and the intermixing of the InAs wetting layers and GaAs spacer layers. The model indicates that the wetting has an alloyed profile, which decreases exponentially from $In_{.99}Ga_{.01}As$ to $In_{.10}Ga_{.90}As$ in 18 Å, or nearly two-thirds of the dot height. In addition, a small amount of In is present throughout the GaAs region. This diffusion of InAs in GaAs has been observed by other researchers in thin InAs QWs in GaAs.[48–50] The diffusion is the result of surface segregation of In. Like the InAs island formation, the surface segregation is a mechanism by which the lattice-mismatched strain can be reduced. After InAs has been deposited on the GaAs surface and GaAs deposition again begins, the In atoms switch sites with surface Ga adatoms. This strain-driven diffusion of In to the surface reduces the total system energy because the surface In is not completely constrained by the GaAs lattice. The complete sequence of In diffusion occurs as follows. Through the initial deposition of less than 2 ML of equivalently planar InAs, the InAs surface remains generally planar, except for some surface roughness that may be present to reduce strain. With the addition of a small amount of In to the surface, the surface transforms from planar to 3-D islands. However, through XRD, we have shown that the wetting layer region remains. If the InAs deposition is terminated before plastic relaxation and the surface is covered with GaAs, In from the planar region surface segregates, leading to an alloyed region between the dots. A consequence of this segregation is that the InAs quantum dots are now less laterally confined because the vertically diffused In decreases the lateral confining potential.

In contrast to the In diffusion present in spaces between the dots, the In in the dot regions is remarkably stable. This can be seen in Fig. 22b, where an STM empty state tunneling image of three vertically aligned dots is shown. The shape of these dots is pancake-like and is the typical shape we observe in all samples. Although the bottom interface of each dot is abrupt, the top interface of the dot is also quite sharp; no isolated In atoms are observed in the vertical regions between the dots, as are seen in the spacer regions away from the dots. In addition, no In concentration gradient is observed to go away from the dots into the wetting layer regions (in the in-plane direction), further indicating the stability of the In in the dots.

We believe that the InAs in the dot regions is resistant to lateral and vertical diffusion because this InAs exists in regions where the lattice constant remains slightly dilated toward the InAs lattice constant. In the wetting layer regions, the InAs is strained to the GaAs lattice constant, and the full effect of strain-enhanced diffusion is present. In fact, if the GaAs regions surrounding the InAs dots have the GaAs lattice constant, then the small difference in lattice constant between the two regions inhibits InAs diffusion away from the dots (Eq. 10).

Increased Structural Uniformity

It has been theoretically and experimentally shown[18,41] that the vertical column structure increases the upper layer dot size uniformity and nearest-neighbor spatial uniformity. The increased uniformity is the result of stress fields from the subsurface dot layers affecting the adatom diffusion on the new surface layer. Suppose that the initial

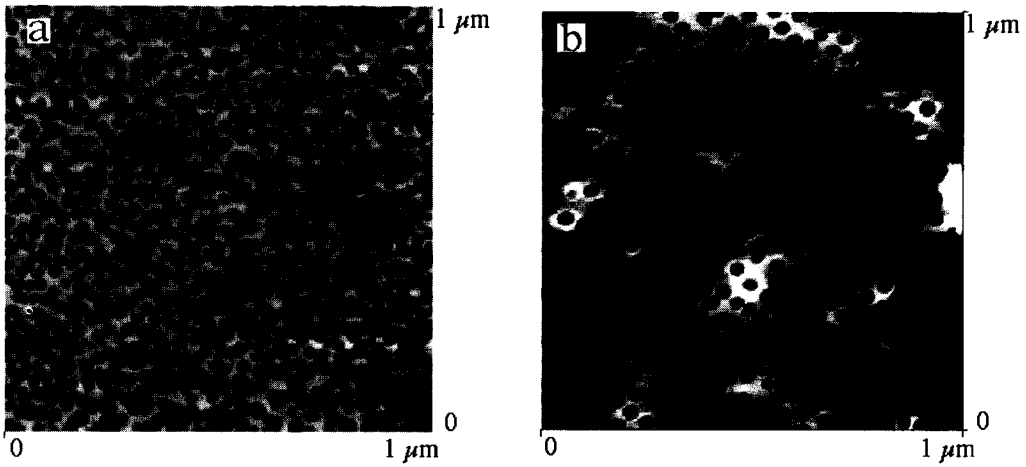


Fig. 24 AFM image of (a) the top surface of a single layer of InAs dots, and (b) of the top surface after 20 layers of dots. (*View this art in color at www.dekker.com.*)

dot density is large compared with the adatom surface diffusion length, and the strain fields from adjacent columns overlap. Because the nucleation of an InAs island occurs at a strain maximum, overlapping columns produce maxima that are closer together. With increased layering of islands, the strain maxima become closer together and columns will eventually be eliminated. If the initial dot density is sparse compared with the adatom surface diffusion length and the island strain distribution, new dots will nucleate between existing columns. Overall, some equilibrium density of islands will result. In this state, the strain field should be periodic over the growth surface; the spacing between strain maxima will be determined by strain, overlayer thickness, and surface diffusion. The nearest-neighbor distance between nucleating dots will become constant, resulting in uniform nutrient fields around each dot and thus a more uniform dot size distribution. In our work, we assume that the ideal equilibrium surface of a single dot layer is a dense array of dots in which the dots form from strain-induced surface roughness. Therefore the formation of columns should decrease the dot density in the uppermost dot layer.

A decrease in island density and an increased island uniformity are immediately apparent when comparing the bare InAs dot surfaces after a single layer of dots and after 20 layers of dots have been deposited. This is shown in the AFM images of Fig. 24. A statistical evaluation of the AFM images in Fig. 24 is found in Figs. 25 and 26. In Fig. 25, a comparison of the variation in height in the single-layer and 20-layer samples is shown. The average height increases from 43 Å in the single sample to 65 Å in the 20-layer sample. The full width at half maximum (FWHM) of the dot height is reduced from 17 to 10 Å between the single-layer and 20-layer samples, even

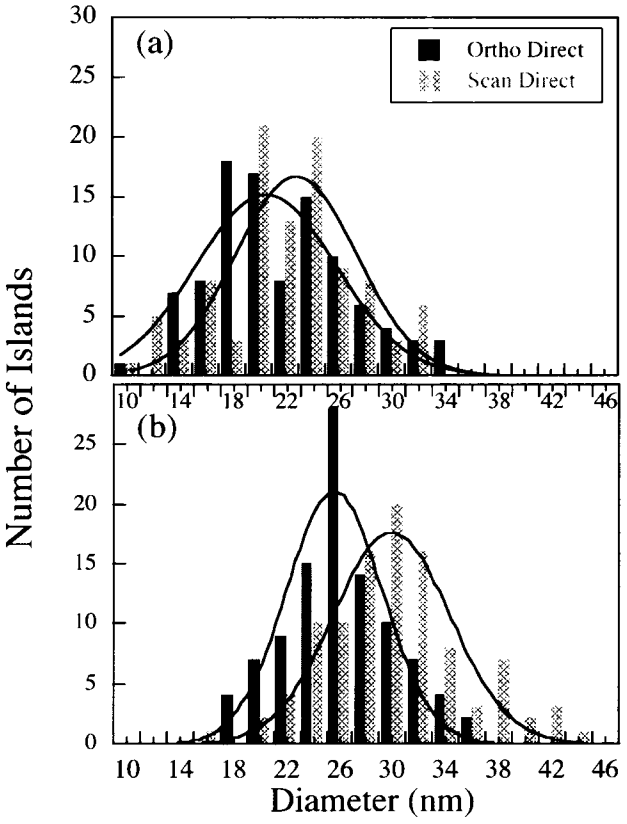


Fig. 25 In-plane dot diameter variations determined from AFM, where (a) is for a single layer of InAs dots and (b) is after 20 layers of InAs dots. (*View this art in color at www.dekker.com.*)

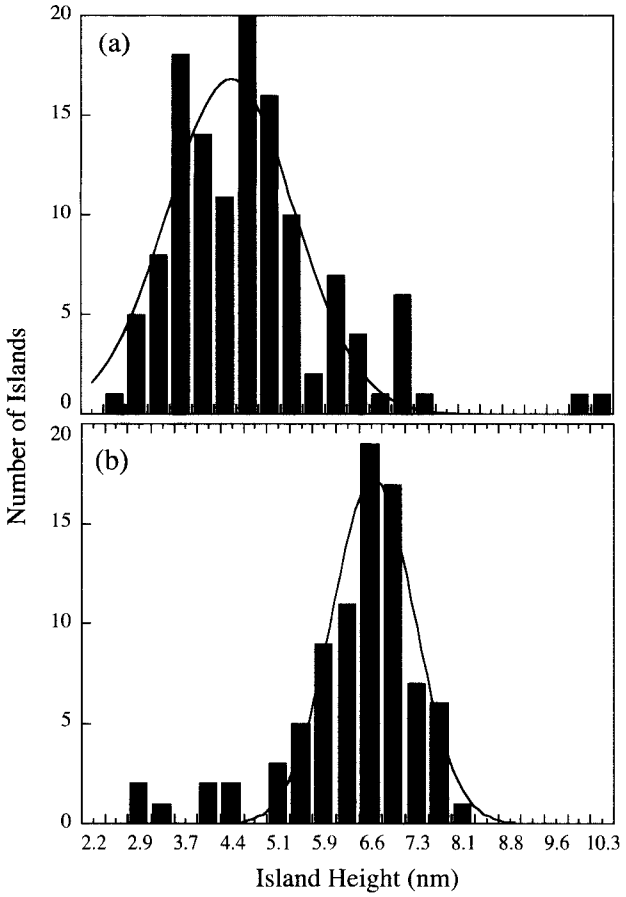


Fig. 26 Height variations determined from AFM (a) for a single layer of dots and (b) after 20 layers of dots. (*View this art in color at www.dekker.com.*)

though the average dot height has increased in the 20-layer sample. In Fig. 26, a comparison of the variation in the in-plane dot diameter between the two samples is shown. The in-plane area is not circular but elliptical, and the two differently shaded regions in Fig. 26 represent the orthogonal directions with maximum and minimum valves. (We cannot observe the facet planes in our AFM imaging, and it is assumed that the island base is rectangular or parallelogram.) As for the average dot height, the in-plane diameter has increased after 20 dot layers. The average in-plane diameter of the single dot layer is 218 Å, whereas the average in-plane diameter after 20 dot layers is 273 Å. The distribution of the in-plane diameter has also narrowed: The FWHM of the diameter distribution is reduced from 50 Å (single layer) to 41 Å (20 dot layers). There is a small but observable difference between the height increase and the diameter increase in the 20-layer dot distribution. The average height increases more than the average diameter, so that the ratio of height to in-plane radius changes from 0.4 to 0.48 after 20 layers (i.e., these dots are more hemispherical). The increase in dot height and diameter is in agreement with the reduced island density observed in Fig. 24. Because the dot density is reduced and the wetting layer thickness is unchanged (Fig. 23a), the dot size must increase. The reduction in the size variation in height and diameter is because of the more uniform local dot environment resulting from stacking the dot layers. Because the strain field in the dot layers becomes more periodic with each dot layer, dots are nucleated with more uniform nearest-neighbor distances, and hence more uniform nutrient fields. These more uniform surroundings lead to more uniform dot sizes.

Cross-sectional STM reveals the way in which the dot columns are eliminated. Empty states images of 20 dot layers are shown in Fig. 27. In Fig. 27a, three columns that

Fig. 27 Cross-sectional STM images in two regions of a 20-layer dot column structure. On the left (a), column uniformity is maintained except for a small truncated column. On the right (b), two columns merge, and then the resulting column merges with a third column. (Courtesy of W. Wu and J. R. Tucker.)

are not terminated are shown, and indicate the well-aligned dot column structure. In Fig. 27b, several columns are eliminated through the dot layering process. Two types of column elimination are present, yet it is not clear whether these two types of elimination are indeed distinct. The first process is by simple extinction: The column abruptly ends. Such a column is observed in Fig. 27b on the left. The second type of column elimination is observed on the right: Two columns merge after only two dot layers, whereas still another column dramatically bends toward this stable column until it eventually merges and is eliminated. Because we only probe in one cross-sectional plane, we cannot be sure that the column that abruptly ends does not bend inward away from the surface; however, the abruptness in which the column ends suggests that this is not the case.

If one column exerts more of an effect on the resulting strain field, then the position of the new dot will be closer to this column. Such a description will account for the abrupt type of column elimination we observe. However, the slow bending of columns is a more dynamic process. In this case, we believe that overlapping strain fields from columns produce a situation in which distinct minima in misfit exist near each column, but that an overall lattice change between the columns shifts one or more of these minima toward the region between columns. It is unlikely that the observed bending is caused by competing nutrient fields because this situation would drive columns away from each other, and this is not observed.

Increased Spatial Uniformity

When a large number of island layers are deposited, the strain created by adjacent columns interacts to create a surface strain distribution. We have just shown how this imposed strain distribution leads to structural ordering of the quantum dot ensemble on the surface. Similarly, the strain distribution becomes a template that can lead to spatial ordering of the new island layer. In investigating the limits of this ordering, we have determined that, under

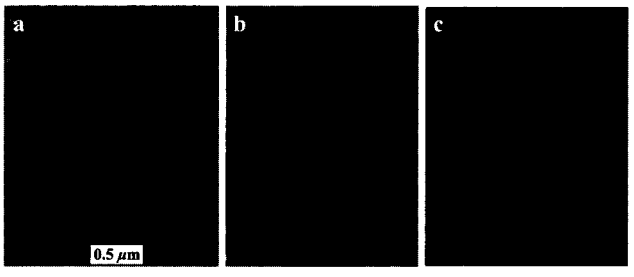


Fig. 28 AFM images of after (a) 30 layers of InAs dots (T_g=500°C), (b) 75 layers of InAs dots (T_g=525°C), and (c) a single layer of dots grown to enhance surface step bunching.

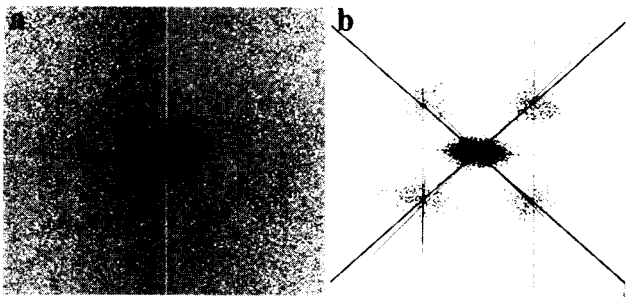


Fig. 29 Fourier transform of the AFM images of the surface of two InAs QD samples: (a) single-QD layer and (b) 30-QD layer.

various crystal growth conditions, a surface unit cell of islands is formed on the upper layer of the superlattice. Although some translational properties of the unit cell are observed, providing initial evidence of a 2-D surface lattice of quantum dots, a well-ordered lattice is not found. We demonstrate that this occurs because in this material system, with a (100) growth surface, surface diffusion effects and bulk strain effects compete in a manner that breaks complete translational symmetry. Adjusting growth parameters can diminish this effect.

An example of the variation in quantum dot ensembles with variations in spacer layer thickness and number of layers can be seen in Fig. 28. The first two images in Fig. 28 have different subsurface InAs dot layer parameters. In Fig. 28a, the subsurface multidot structure has 30 InAs dot layers, each containing the equivalent of 2.5 ML of planar InAs. The thickness between each wetting layer region is 55 Å, and the growth temperature was 500°C. In Fig. 28b, the subsurface multilayer contains 50 dot layers, the wetting layer spacer is 75 Å, and the growth temperature is 525°C. In "Single-Layer InAs Quantum Dot Ensembles" and our previous work,[36] we showed that the higher the growth temperature is, the larger is the average InAs dot size. Thus the GaAs spacer layer must be increased in the sample grown at a higher temperature so that the growth surface can planarize between dot layers. In Fig. 28a and b, a qualitative alignment along ⟨100⟩ can be seen. Ordering in single-layer samples, if present at all, is along [110] and is because of step bunching.[6] An example of this is shown in Fig. 28c, where the growth parameters where adjusted to maximize step bunching. In contrast to the ⟨100⟩ alignment seen in Fig. 28a and b, the alignment in Fig. 28c is in the vertical [110] direction. The alignment is caused by anisotropic surface diffusion because of ledge and surface reconstruction.[51] In contrast, the surface alignment of InAs dots in ⟨100⟩ in samples containing subsurface dot multilayers is because of the interaction of the strain fields from the buried dots.

We can begin to see the extent of the spatial ordering present in these structures by making reciprocal space

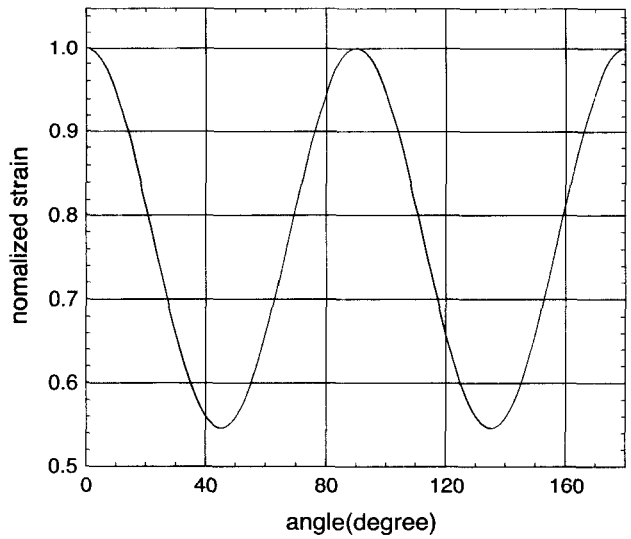


Fig. 30 The normalized angular variation in strain on the GaAs surface caused by an island below the surface. Zero degrees is taken in the [100] direction.

surface maps of AFM images. The Fourier transforms of the AFM images for the single-layer and the 30-layer samples shown in Fig. 28 are shown in Fig. 29. In the Fourier space image of the single island layer sample, only a centered peak is present, indicating there is no regular nearest-neighbor distance between islands. In the 30-layer sample, there is both regular nearest-neighbor distance and orientation. The nearest-neighbor directions are only approximately the $\langle 100 \rangle$ directions, as shown by the thin lines added in the vertical and diagonal directions. Our explanation for this is that, although the strain distribution on the surface drives the unit cell of dots to have a nearest-neighbor direction in $\langle 100 \rangle$, the surface diffusion has a different preferential orientation. The anisotropic surface diffusivity in a sense competes with subsurface strain and pushes the nearest-neighbor direction slightly away from $\langle 100 \rangle$. We will show later in this chapter that the situation is even more interesting because the second nearest-neighbor directions from the strain distribution will be shown to be $\langle 110 \rangle$, but because of the anisotropic surface diffusion, these unit cell directions are no longer degenerate.

In the case of an isotropic material, the region above a subsurface island is strained and becomes an advantageous site for island formation because the dilated lattice constant is closer to the InAs lattice then the surrounding surface. The strain interaction of two subsurface islands has been calculated for the isotropic case.[52] We use one island as an infinitesimal "test island" in analogy to the test charge in electrostatics, with the convention that it induces a negative strain in the crystal. The strain distribution on the surface from an island submerged a distance L from the surface is:[18]

$$\epsilon(x) = C\frac{1}{(x^2 + L^2)^{\frac{3}{2}}}\left(1 - \frac{3L^2}{(x^2 + L^2)}\right) \tag{13}$$

where C is constant. If two islands are in close proximity, the resulting surface strain distribution is the superposition of the strain distributions of the individual islands. On the 2-D surface, the possible first nearest-neighbor island positions fall on a circle. If the surface diffusion rate is adequate, the highest density of islands (a hexagonal cell) will result. If enough layers are deposited, different surface regions will coalesce into a close-packed 2-D hexagonal arrangement.

What is interesting to us is the effect of the crystal symmetry on this result. Thus for the moment, we ignore the surface diffusion effect and determine the nearest-neighbor arrangement from crystal symmetry. The nearest-neighbor ordering observed in Fig. 29 can be explained by the anisotropic properties of the GaAs zinc-blend crystal structure. The $\langle 100 \rangle$ directions are elastically soft[53] (i.e., they have relatively small elastic moduli). Thus it is expected that nearest-neighbors should be in these directions. We investigate the effect of the anisotropic elasticity tensor on the orientation of the surface ordering by calculating the surface strain from a single buried island as a function of crystallographical orientation, and then assuming that the total strain from a buried layer of many islands is the linear superposition of the strain from individual islands. We do this by using a known isotropic solution[54] for an isolated subsurface island and adopting it to the anisotropic case. A more accurate but still approximate solution can be developed using the elastic stress tensor of a cubic crystal. However, approximations must still be made and, for our purposes, this simple solution is adequate. We use a Green's function approach for a buried elliptical island where the island dimensions have been determined from the cross-sectional STM discussed in "Increased Structural Uniformity." Following Hooke's law, the stress σ and strain are related by the elastic tensor c_{ijkl} by $\sigma = c_{ijkl}\epsilon$. For rotation about the surface normal, we assume that the inverse of the elastic tensor (denoted by s_{ijkl}) is a slowly varying function of the orientation in GaAs, so that the already known isotropic Green's functions can be applied. The displacement is related to stress σ^* induced by the buried InAs island by $c_{ijkl}\mathbf{u}_{ij,kl} = \sigma_{ij}$, where ij,kl indicates partial differentiation by the cyclical indices k and l. Then:

$$u_i(x) = -c_{jlmn}\varepsilon * \int_{\Omega} G_{ij,l}(\mathbf{x} - \mathbf{x}')\mathrm{d}x' \tag{14}$$

is a solution for the displacement. An image island above the surface is used to force the surface tractions to

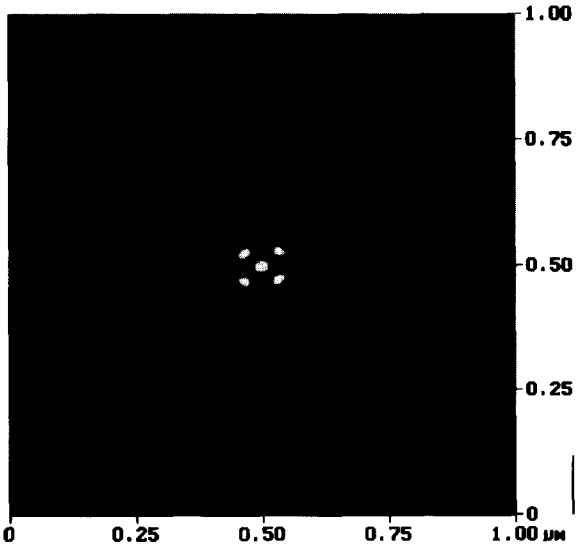


Fig. 31 2-D autocovariance of the AFM image in Fig. 28a.

zero. The strain is related to the displacement by $\epsilon_{ij}=1/2(u_{i,j}+u_{j,i})$ and leads to surface strain as a function of position x and crystallographical orientation θ:

$$\epsilon_{ij}(x,\theta) = -\frac{\epsilon* \; s_{11}(\theta)+s_{12}(\theta)}{4\pi \; s_{11}(\theta)-s_{12}(\theta)}\left(\int_{\Omega_1}\left(\frac{1}{R_1}\right)_{ij}\mathrm{d}x' + \left(3+\frac{4s_{12}(\theta)}{s_{11}(\theta)}\right)\int_{\Omega_2}\left(\frac{1}{R_2}\right)_{ij}\mathrm{d}x'\right) \quad (15)$$

The subsurface island is region Ω_1, whereas the image island is region Ω_2. R_1 is $|x-x'|$, where x' is on Ω_1, and R_2 is the distance between a point in Ω_2 and x. The θ dependence of the above equation can be evaluated by transformation of the stiffness constants $s_{ijkl}(\theta)$ into a rotating coordinate system on the surface and centered above an island. This transformation is $s_{ij}=\sum_{m,n}s_{mn}q_{mi}q_{nj}$, where q_{ij} is the transformation coefficient and is of second order in the directional cosines.

Fig. 30 shows the variation of the surface strain as a function of crystallographical orientation. The maximum strain is in the ⟨100⟩ directions. Because InAs island nucleation will occur preferentially where the GaAs crystal is biaxially expanded, this direction becomes the nearest-neighbor direction.

In the Fourier-transformed AFM images, we would like to investigate higher-order modes related to the second nearest-neighbors. Therefore we have examined the autocovariance of the images. In Fig. 31, the autocovariance of the AFM image from Fig. 28a is shown. This sample contains 30 InAs dot layers. The autocovariance of a function f is defined as:

$$\mathbf{R}_{ff}(\mu,\nu) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} f(x,y)f(x+\mu,y+\nu)\mathrm{d}x\mathrm{d}y \quad (16)$$

Here f is a functional representation of the real space image in the x–y plane, μ and ν are the x and y translations, and T is a spatial limit. We now investigate primary unit cell directions by taking slices through this transformation. Fig. 32 shows autocovariance data from two samples in the ⟨110⟩ and ⟨100⟩ directions. Peaks in the figure indicate a translational correlation. In the single-layer

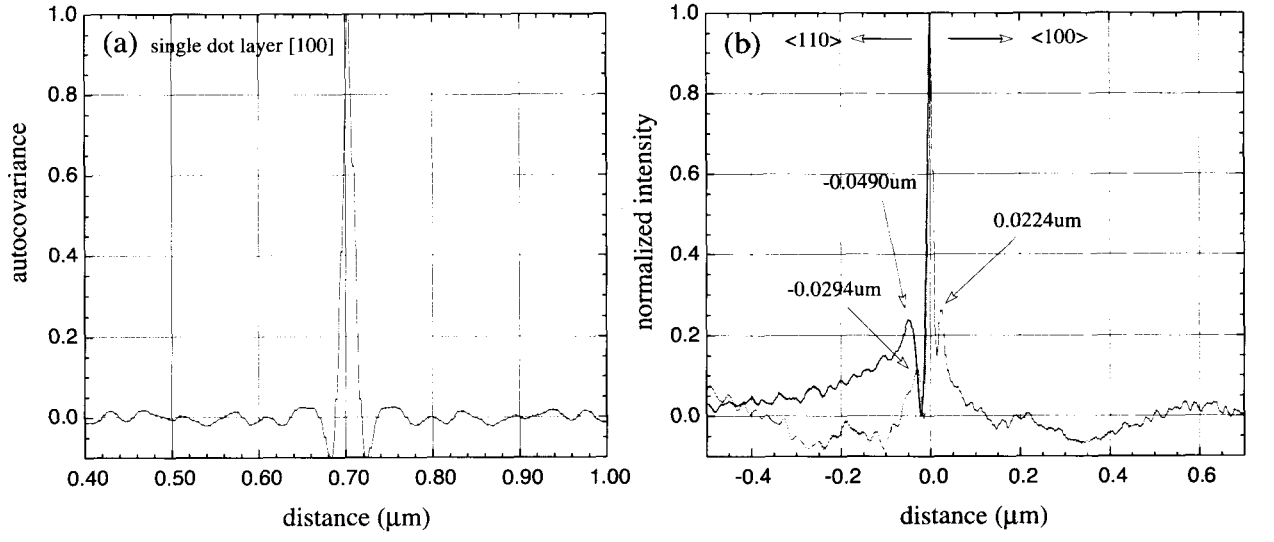


Fig. 32 Autocovariance of AFM images in the ⟨110⟩ and [100] directions: (a) a single dot layer sample, and (b) a 30-layer dot sample. The 30-layer sample has 55 Å spacer layers and was fabricated at 500°C. (*View this art in color at www.dekker.com.*)

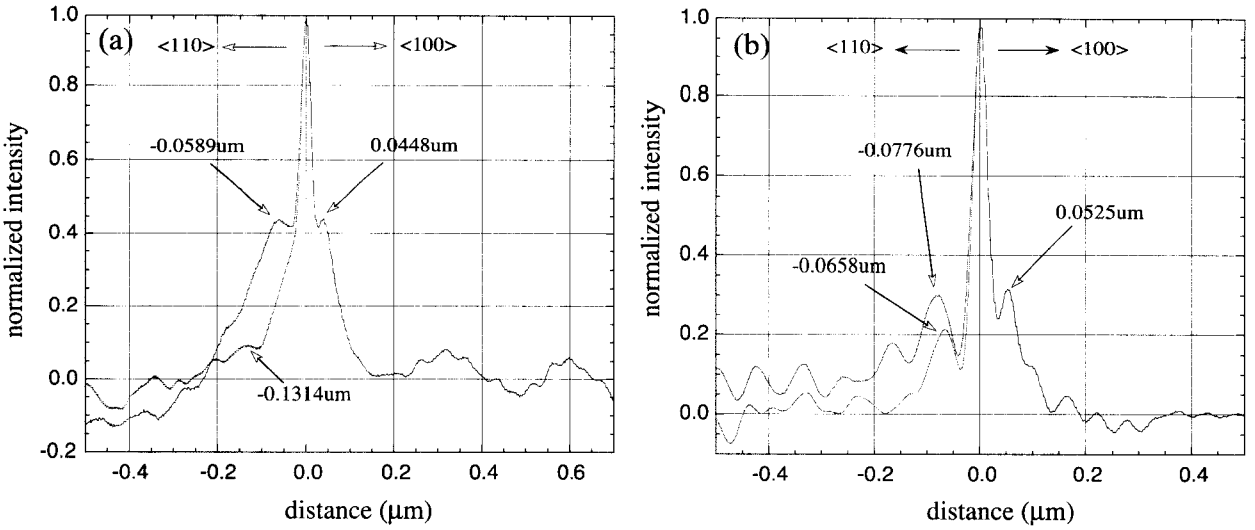


Fig. 33 Autocovariance of AFM images in the ⟨110⟩ and [100] directions for two structures fabricated at 525°C: (a) 50 layers of InAs dots were used in the subsurface structure, whereas in (b), 75 layers were used. (*View this art in color at www.dekker.com.*)

case, shown in Fig. 32a, there is only a central peak, indicating no correlation between islands. In Fig. 32b, we show autocovariance sections of a sample grown at 500°C with 30 InAs dot layers, separated by 55 Å GaAs spacer layers. Here the single satellite peaks in the ⟨110⟩ and ⟨100⟩ directions near the central peak indicate the InAs quantum dot unit cell; once again, ⟨100⟩ is only approximate. From the figure, the first nearest-neighbor directions are in the predominantly ⟨100⟩ directions, whereas the second nearest-neighbor direction is [$\bar{1}$10] and the third nearest-neighbor direction is [110]. Because the subsurface strain should create four nearest-neighbors in ⟨100⟩, we believe that both the deviation from this direction and the nondegeneracy of the orthogonal second and third nearest-neighbors result from the anisotropic surface diffusion. The longer-wavelength, lower-intensity oscillations in Fig. 32b are the result of large-scale surface roughness, possibly because of steps on the surface. On top of this low-frequency roughness, some periodic intensity fluctuations can be seen, especially in one of the ⟨110⟩ directions. This periodicity matches the fundamental unit cell distance and shows the translational symmetry in these directions.

To optimize the unit cell periodicity, we can adjust the crystal growth parameters. Fig. 33 shows the autocovariance of images from two samples grown at 525°C. In Fig. 33a, 50 layers of InAs dots have been used, whereas in Fig. 33b, the number of InAs dot layers has been increased to 75. In both cases, the spacer layer thickness is 100 Å. Compared with the earlier samples grown at 500°C, this spacer layer thickness has been increased because of the increased dot size at increased growth temperatures. In Fig. 33a, nearest-neighbor periodicity can be seen in both the ⟨110⟩ and [100] directions. The nearest-neighbor distance has grown from 225 to 525 Å because of the increasing dot size with increasing growth temperature.[36] In Fig. 33b, the autocovariance shows a much more pronounced periodicity in the first, second, and third nearest-neighbor directions. We now see clear evidence of the emergence of a surface lattice of QDs. The condition that improves the surface lattice structure is primarily increased growth temperature.

We can combine the information from Figs. 32 and 33 to begin to understand why the quantum dots form only a weakly organized lattice. The lattice is weak because the ratio of the first, second, and third nearest-neighbor distances does not allow for the complete filling of the surface. The anisotropic surface strain distribution creates a template for a cubic lattice of InAs dots. This process would create first nearest-neighbor sites in the ⟨100⟩ directions. However, the surface diffusion of In or Ga adatoms is also anisotropic because of surface reconstruction and ledge dynamics. Because the surface diffusion is larger in [1$\bar{1}$0],[51] the linear dot density in this direction should be largest. When these two processes are combined, the degeneracy in the ⟨110⟩ directions is lifted and the second nearest-neighbor sites in [1$\bar{1}$0] become closer than in [110]. In addition, the first nearest-neighbor site is rotated toward [110] because of the anisotropic surface diffusion.

The angular difference from the ideal first nearest-neighbor position (⟨100⟩) and the experimentally determined position is approximately 3.5°. This small difference from the ideal subsurface strain-driven case, where the first nearest-neighbor orientation should be ⟨100⟩, shows that the surface diffusion does not dominate the island ordering. However, it is enough to inhibit the

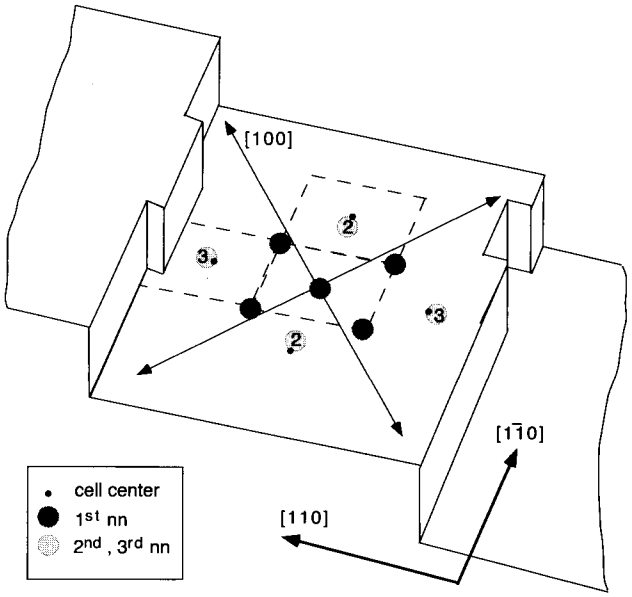


Fig. 34 The unit cell is shown on a ledged surface. The unit cell is shown as a broken line whose center is translated to the small black dots. Experimentally determined first, second, and third nearest-neighbor sites are indicated by larger dots. The translated unit cell does not align with the second and third nearest-neighbors.

formation of a well-ordered surface lattice. A cartoon of the quantum dot unit cell is shown in Fig. 34. The large dark circles represent the first nearest-neighbors, and the slight misalignment from ⟨100⟩ can be seen. The second and third nearest-neighbors are represented by the large, lighter-colored circles. The anisotropy in the surface diffusion length,[51] attachment kinetics, and nucleation in orthogonal ⟨110⟩ directions makes the nearest-neighbor distance in [1$\bar{1}$0] shorter than the nearest-neighbor distance in [110]. Finally, because of the anisotropy on the surface, the surface QD unit cell cannot be translated to fill the surface—no lattice is formed. This is seen in Fig. 34, where the small black dots represent the center of the translated unit cell. These smaller black dots do not fall on the second and third nearest-neighbor sites.

As the growth temperature increases, the anisotropy is reduced; this is shown in Fig. 35. As the growth temperature is increased, the difference in length between the second and third neighbors is reduced and the deviation from the ideal ratios between the first, second, and third nearest-neighbors is reduced. From Fig. 32 (510°C) and Fig. 33b (525°C), the ratio between the second and third nearest-neighbors is reduced from 1.70 to 1.15. The reduction is anisotropy shows that the effect of surface diffusion is reduced. If the second and third nearest-neighbors are used to calculate the first nearest-neighbor position based on a simple rectangular lattice, the deviation from the ideal position drops from 0.22 to only 0.035 as the temperature is increased from 510°C to 525°C. Thus as the growth temperature increases, a more ideal simple rectangular unit cell can be formed. This more ideal unit cell is more easily translated on the surface to form a lattice. We believe that the increased growth temperature reduces the anisotropy in the surface diffusion. Although an increased growth temperature would

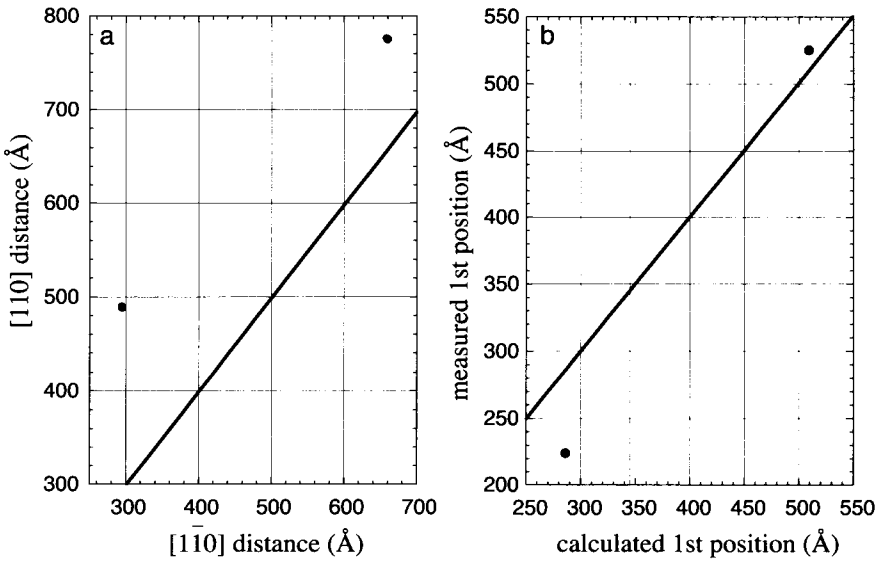


Fig. 35 The variation from an ideal rectangular unit cell is shown for samples grown at 510°C and 525°C. In (a), the [110] to [1$\bar{1}$0] nearest-neighbor distances are compared. The solid line indicates an ideal rectangular cell. In (b), the first nearest-neighbor distance is calculated based on the second and third nearest-neighbor distances, and compared to the measured values. (*View this art in color at www.dekker.com.*)

create larger terrace lengths in an unstrained GaAs surface, this may not be the case in the strained surface. Further work is necessary to confirm these results and develop a better understanding of the interplay between the subsurface strain and the surface diffusion. However, from this work, the spatial ordering from a subsurface superlattice structure of dot layers is improved with increased growth temperature.

Finally, it is important to note that the competition between surface diffusion and subsurface strain shown here is a manifestation of our particular crystal system and substrate orientation. Other systems with different surface reconstructions and ledge dynamics, or other vicinal or crystallographical surfaces will certainly act differently.

CONCLUSION

We began this chapter by highlighting certain aspects of strain-induced quantum dot formation. These dots are formed in dense arrays. Although the density and uniformity of these ensembles can be controlled by adjusting the growth parameters, inhomogeneous broadening continues to mask the atomlike nature of the quantum dot states. Two features of this system were discussed. First, the luminescence linewidth was compared with the dot structure. It was shown that, although the linewidth is reduced in samples made at lower growth temperatures, the dots are less distinct. The decreasing linewidth is the result of reduced localization as the wavefunctions of individual dots in the ensemble begin to extend into adjacent dots. Second, the density of the dot ensemble was shown to be a function of the V/III flux ratio. In some applications, a high density is important, and we show how to maximize the density and implications of the process on dot formation.

The main purpose of this chapter is to develop the idea of using stacks of InAs quantum dot layers. These structures can be used in two ways. First, when only a small number of dot layers are stacked, columns of dots form electronically coupled structures. In analogy to coupled QWs, changing the number of dot layers and the spacer layer thickness between them can tune the coupling. As a simple test of the structures, LEDs were fabricated, and it was shown that luminescence from the wetting layer and higher dot states could be suppressed using these column structures. Second, when a large number of dot layers are used, the surface size and spatial distributions can be improved. However, the experimental improvements are limited in our system because of competition between subsurface strain and surface diffusion. The way in which these two processes interact is clearly evident in the AFM images. Higher growth temperatures reduce the anisotropic surface diffusion effect, leading to an extended lattice. This lattice is present, but is not perfect, and perhaps other substrate orientations or material systems may offer improvements.

This work is just the beginning of efforts to create ensembles of atomlike quantum dots using epitaxial crystal growth. Although GaAs and InAs are used in these experiments, this work is not limited to III–V materials or even semiconductors. If well-ordered arrays of quantum dots can be created, promising device and fundamental physics opportunities await. In the device world, regular arrays of quantum dots allow for lithographical alignment and possibilities of logic and memory devices. As well, many of the experiments in atomic physics can be repeated using quantum dots. An interesting extension of this work would be the combination of the two aspects of the vertical column structures: the fabrication of structures in which a subsurface island superlattice of one material is used to make a surface strain distribution that increases the ordering in the surface layers, whereas a small number of quantum dots of a different material are grouped together in short columns as a coupled active region.

ACKNOWLEDGMENTS

This work was only made possible through collaborations with and support from many people. The STM images were performed by W. Wu in J. R. Tucker's laboratory at the University of Illinois at Urbana-Champaign. TEM images where made with A. F. Marshall (Center for Materials Research at Stanford University). AFM and autocovariance data were obtained and analyzed with the help of S. Komarov (Photonics and Electronics Laboratory at Stanford University). G. S. would like to thank Y. Yamamoto for continued support. We would like to thank M. Pelton for critical discussions and helpful comments. This work was supported by the DARPA and ONR through contract nos. N00014-93-1 and N00014-97-1, and by the JST through the ERATO Quantum Fluctuation Project.

REFERENCES

1. Matsumoto, K.; Ishii, M.; Segawa, K.; Oka, Y.; Vartanian, B.J.; Harris, J.S., Jr. Appl. Phys. Lett. **1996**, *68*, 34.
2. Tarucha, S.; Austing, D.G.; Honda, T.; Hage, R.; Kouwenhoven, L.P. Jpn. J. Appl. Phys. **1997**, *36*, 3917.
3. Yano, K.; Ishii, T.; Hashimoto, T.; Kobayashi, T.; Murai, F.; Seki, K. Appl. Phys. Lett. **1995**, *67*, 828.
4. Bockelmann, U.; Heller, W.; Filoramo, A.; Roussignol, Ph. Phys. Rev., B **1997**, *55*, 4456.
5. Ipatova, I.P.; Malyshkin, V.G.; Shchukin, V.A. J. Appl. Phys. **1993**, *74*, 7198.

6. Tersoff, J.; Phang, Y.H.; Zhang, Z.; Lagally, M.G. Phys. Rev. Lett. **1995**, *75*, 2730.
7. Lijima, S. Nature **1991**, *354*, 56.
8. Tanford, C. Science **1978**, *200*, 1012.
9. Tabuchi, M.; Noda, S.; Sasaki, A. *Science and Technology of Mesoscopic Structures*; Springer-Verlag: Tokyo, Japan, 1992; 375.
10. Leonard, D.; Krishnamurthy, M.; Reaves, C.M.; Denbaars, S.P.; Petroff, P.M. Appl. Phys. Lett. **1993**, *63*, 3203.
11. Medeiros-Ribeiro, G.; Lenard, D.; Petroff, P.M. Appl. Phys. Lett. **1995**, *66*, 1767.
12. Heitz, R.; Grundmann, M.; Ledentsov, N.N.; Eckey, L.; Veit, M.; Bimberg, D.; Ustinov, V.M.; Egorov, A.Yu.; Zhukov, A.E.; Kop'ev, P.S.; Alferov, Sh.I. Appl. Phys. Lett. **1996**, *68*, 361.
13. Hessman, D.; Castrillo, P.; Pistol, M.E.; Samuelson, L. Appl. Phys. Lett. **1996**, *69*, 749.
14. Heinrichsdorff, F.; Mao, M.H.; Kirstaedter, N.; Krost, A.; Bimberg, D.; Kosogov, A.O.; Werner, P. Appl. Phys. Lett. **1997**, *71*, 22.
15. Solomon, G.S.; Trezza, J.A.; Marshall, A.F.; Harris, J.S., Jr. Phys. Rev. Lett. **1996**, *76*, 952.
16. Solomon, G.S.; Trezza, J.A.; Harris, J.S., Jr. J. Vac. Sci. Technol., B **1996**, *14*, 2208.
17. Xie, Q.; Madhukar, A.; Chen, P.; Kobayashi, N.P. Phys. Rev. Lett. **1995**, *75*, 2542.
18. Tersoff, J.; Teichert, C.; Lagally, M.G. Phys. Rev. Lett. **1996**, *76*, 1675.
19. Tiller, W.A. *The Science of Crystallization: Microscopic Interfacial Phenomena*; Cambridge University Press: Cambridge, England, 1991; 172.
20. Gao, H. Int. J. Solids Struct. **1990**, *28*, 703.
21. Freund, L.B.; Jonsdottir, F. J. Mech. Phys. Solids **1993**, *41*, 1245.
22. Orr, B.G.; Kessler, D.; Synder, C.W.; Sander, L. Europhys. Lett. **1992**, *19*, 33.
23. Vanderbilt, D.; Wickham, L.K. Mater. Res. Soc. Symp. Proc. **1991**, *202*, 555.
24. Matthews, J.W.; Blakeslee, A.E. J. Cryst. Growth **1974**, *27*, 118.
25. People, R.; Bean, J.C. Appl. Phys. Lett. **1985**, *47*, 322.
26. Stranski, I.N.; Krastanow, L. Akad. Wiss. Wien Math. Naturwiss. Kl. Abt. 2B Chemie **1937**, *146*, 797. We add this reference for historical context.
27. Mullins, W.W. J. Appl. Phys. **1957**, *28*, 333.
28. Asaro, R.J.; Tiller, W.A. Metall. Trans. **1972**, *3*, 1789.
29. Srolovitz, D.J. Acta Metall. **1989**, *37*, 621.
30. Lee, H.; Lowe-Webb, R.; Yang, W.; Sercel, P.C. Appl. Phys. Lett. **1998**, *72*, 812.
31. Gao, H. J. Mech. Phys. Solids **1991**, *39*, 443.
32. Gao, H. J. Mech. Phys. Solids **1994**, *42*, 741.
33. Freund, L.B.; Johnson, H.T.; Kukta, R.V. Mater. Res. Soc. Symp. Proc. **1995**, 359.
34. Guha, S.; Madhukar, A.; Rajkumar, K.C. Appl. Phys. Lett. **1990**, *57*, 210.
35. Hellman, E.S.; Harris, J.S., Jr. J. Cryst. Growth **1987**, *81*, 38.
36. Solomon, G.S.; Trezza, J.A.; Harris, J.S., Jr. Appl. Phys. Lett. **1995**, *66*, 991.
37. Leonard, D.; Pond, K.; Petroff, P.M. Phys. Rev., B **1994**, *50*, 11687.
38. Lin, X.W.; Liliental-Weber, Z.; Washburn, J.; Weber, E.R. J. Vac. Sci. Technol., B **1994**, *12*, 2562.
39. Marzin, J.Y.; Gerard, J.M.; Izrael, A.; Barrier, D.; Bastard, G. Phys. Rev. Lett. **1994**, *73*, 716.
40. Fafard, S.; Leonard, D.; Merz, J.L.; Petroff, P.M. Appl. Phys. Lett. **1994**, *65*, 1388.
41. Solomon, G.S.; Trezza, J.A.; Harris, J.S., Jr. Appl. Phys. Lett. **1995**, *66*, 3161.
42. Goldstein, L.; Glas, F.; Marzin, J.Y.; Charasse, M.N.; Le Roux, G. Appl. Phys. Lett. **1985**, *47*, 1099.
43. Nabetani, Y.; Ishikawa, T.; Noda, S.; Sasaki, A. J. Appl. Phys. **1994**, *76*, 347.
44. Leonard, D.; Fafard, S.; Pond, K.; Zhang, Y.H.; Merz, J.L.; Petroff, P.M. J. Vac. Sci. Technol., B **1994**, *12*, 2516.
45. Gerard, J.M.; Genin, J.B.; Lefebvre, J.; Moison, J.M.; Lebouche, N.; Barthe, F. J. Cryst. Growth **1995**, *150*, 351.
46. Lott, J.A.; Ledentsov, N.N.; Ustinov, V.M.; Egorov, A.Yu.; Zhukov, A.E.; Kop'ev, P.S.; Alferov, Zh.I.; Bimberg, D. Electron. Lett. **1997**. *33*, 1150.
47. Mirin, R.P.; Ibbetson, J.P.; Nishiad, K.; Gossard, A.C.; Bowers, J.E. Appl. Phys. Lett. **1996**, *67*, 3795.
48. Zheng, J.F.; Walker, J.D.; Salmeron, M.B.; Weber, E.R. Phys. Rev. Lett. **1994**, *72*, 2414.
49. Pfister, M.; Johnson, M.B.; Alvarado, S.F.; Salemink, H.W.M.; Marti, U.; Martin, D.; Morier-Genoud, F.; Reinhart, F.K. Appl. Phys. Lett. **1995**, *67*, 1459.
50. Sauvage-Simkin, M.; Garreau, Y.; Pinchaux, R.; Veron, M.B.; Landesman, J.P.; Nagle, J. Phys. Rev. Lett. **1995**, *75*, 3485.
51. Ohta, K.; Kojima, T.; Nakagawa, T. J. Cryst. Growth **1989**, *95*, 71.
52. Maradudin, A.A.; Wallis, R.F. Surf. Sci. **1980**, *91*, 423.
53. Okada, T.; Weatherly, G.C.; McComb, D.W. J. Appl. Phys. **1997**, *81*, 2185.
54. Mura, T. *Micromechanics of Defects in Solids*; Martinus Nijhoff: Netherlands, 1982; 93.

Quantum Dots: Inelastic Light Scattering from Electronic Excitations

Q

Christian Schüller
Universität Hamburg, Hamburg, Germany

INTRODUCTION

Electrons confined in semiconductor quantum systems are a field of enormous and still growing research interest because they allow, in specially tailored systems, the investigation of fundamental properties, such as many-particle interactions of electrons in reduced dimensions. In this article, we give an overview of experimental and theoretical investigations of electronic excitations in semiconductor quantum dots. Optical spectroscopy techniques, such as far-infrared (FIR) transmission[1–7] and resonant Raman scattering, i.e., inelastic light scattering (ILS),[8–12] are ideal tools to study the spectrum of elementary excitations of these systems. Since the work of Pinczuk et al.[13] on 2-D intersubband excitations in GaAs–AlGaAs quantum wells, it has been known that besides collective spin-density (SDEs) and charge-density excitations (CDEs), one can observe nearly unrenormalized excitations—the so-called single-particle excitations (SPEs)—in ILS experiments. Both SDEs and CDEs are collective excitations; SDEs are affected by exchange interaction while CDEs are affected by the full Coulomb interaction of the electrons. However, the origin of the SPEs, which seem to be unaffected by the particle–particle interaction, has posed a puzzle. Electronic excitations, and also SPEs in particular, could subsequently be observed in lower-dimensional systems, based on modulation-doped GaAs–AlGaAs quantum wells, and also particularly in quantum dots.[8–10] In an experimental work, it was shown that SPEs can be observed in low-dimensional electron systems under conditions of extreme resonance, when the laser energy is close to the fundamental band gap of the structures.[10] Thus SPEs are created in a resonant density-fluctuation scattering process,[14,15] whereas collective SDEs and CDEs ensue from an excitonic third-order scattering process.[10,16,17] Many theories of nonresonant Raman scattering accurately describe the energetic positions of the collective excitations, as well as the wave-vector and magnetic-field dependence of the CDE and SDE.[18–21] However, they fail in predicting the experimentally observed relative strengths of the different modes. Furthermore, the occurrence of SPE cannot be explained within the confines of these theories. It has been known for a long time that valence-band states play a crucial role[22,23] in carrying out a correct treatment of the resonant scattering cross section. Recent theoretical papers on quantum wires[15,24,25] and quantum dots[26] showed that inclusion of the valence-band states indeed significantly changes the intensities of the excitations. During the past decade, self-assembled InAs quantum dots (SAQDs) have also proven to be highly interesting quantum structures, both from a technological as well as from a fundamental physics point of view. They exhibit relatively large quantization energies in the range of about 50 meV. In most experiments reported so far, SAQDs have been investigated by optical spectroscopy, in particular photoluminescence (PL). Nowadays, PL experiments on single dots are well established, which overcome the inhomogeneously broadened linewidths in typical ensemble measurements.[a] It has also been demonstrated that it is possible to charge SAQD with single electrons[28] via the application of external gate structures. So far, there are only two reports in literature about ILS experiments on electronic excitations in InGaAs SAQD.[29,30] In this contribution, ILS experiments on collective CDEs in InAs SAQD with tunable electron numbers N are described. In these experiments, N can be controlled at $N = 1$–6.

The article is organized as follows. In the section "Characteristics of Quantum Dots and Experimental Details," we describe the quantum-dot structures under investigation: modulation-doped GaAs–AlGaAs quantum dots and InAs SAQD, and give a brief discussion of the electronic structure and the excitations of these systems. Furthermore, the experimental realization of ILS is outlined in this section. In the section "Scattering Mechanisms," a summarized description of the scattering mechanisms, which lead to the creation of electronic excitations in quantum dots, is provided. In the section "Experiments on GaAs–AlGaAs Deep-Etched Quantum Dots," we start to discuss experiments on modulation-doped GaAs–AlGaAs quantum dots. There, basics such as parity selection rules are elaborated. The section "Experiments on InAs Self-Assembled Quantum Dots" summarizes recent experiments on InAs SAQD, which contain

[a]For a recent review, see Ref. [27].

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013811

only a small number of electrons and which can be regarded as artificial atoms.

CHARACTERISTICS OF QUANTUM DOTS AND EXPERIMENTAL DETAILS

Preparation of Quantum Dots

The vast majority of ILS experiments on electronic excitations in quantum dots have been performed on modulation-doped GaAs–AlGaAs structures. Fig. 1a shows a scanning electron micrograph of deep-etched quantum dots. Typical sizes of these structures are on the order of several hundred nanometers. In many cases, 2-D electron systems, realized in modulation-doped single quantum wells, are used as a starting material. In Fig. 1a, the location of such a 25-nm-wide quantum well is indicated by thick dashed lines. By a reactive-ion etching process, the pillar-shaped quantum dots were defined. Typically, such deep-etched structures contain electron numbers on the order of several hundreds. An alternative and very effective way to produce quasi zero-dimensional semiconductor structures is the so-called self-assembled growth of InAs quantum dots. Here the lattice mismatch between GaAs, which is used as the starting material, and InAs is exploited. The minimization of strain energy leads to the formation of small InAs islands, if InAs is grown on a GaAs surface. In contrast to the etched structures, the typical lateral sizes of these islands are on the order of only a few tens of nanometers. Therefore they exhibit much larger quantization energies. Fig. 1b displays a transmission electron micrograph of an InAs quantum dot, embedded in GaAs. The boundaries of the InAs dot are indicated by a white dashed line. It has been shown that electrons can be filled into these quantum boxes either by modulation doping[29] or by application of external gates.[28] With the application of gates, it is even possible to fill in single electrons in a very controlled way. Fig. 1c shows a schematic picture of the band structure of an InAs SAQD sample, as used in Ref. [30]. There, an inverted modulation-doped AlGaAs–GaAs structure serves as a back contact. By application of a voltage between the back contact and a metallic front gate, the dots can be charged with single electrons, which then tunnel from the two-dimensional electron system (2DES) to the dots.

Electronic Ground State and Excitations

As will be explained below, in the etched GaAs dots, as well as in the InAs SAQD,[31] the lateral external confining potential is to a good approximation parabolic. In the deep-etched samples, the potential is determined by the homogeneously distributed ionized remote donors in the AlGaAs barrier and negatively charged surface states. By analytical calculations, one can show that this lateral potential, which acts on a test electron within the structure (*external* potential), is in x and y directions (lateral directions) in good approximation parabolic.[32]

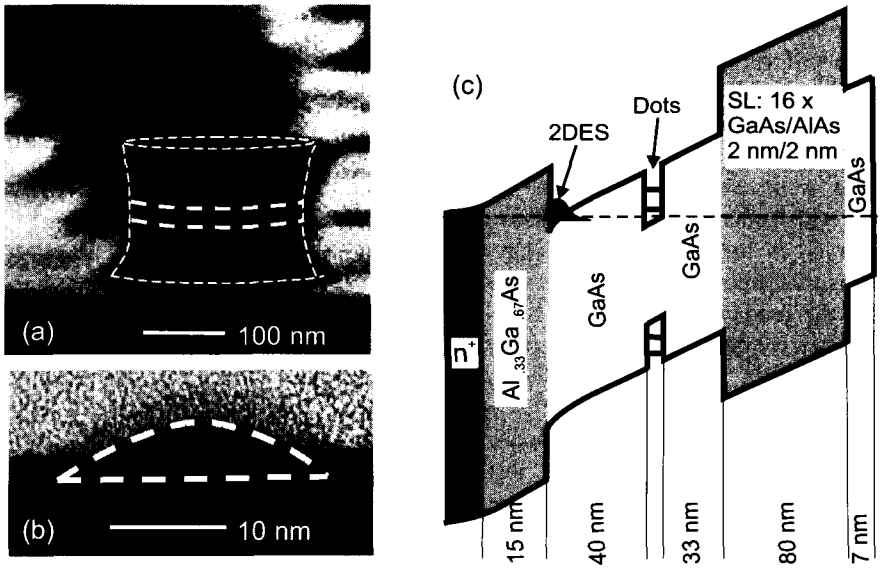


Fig. 1 (a) Scanning electron micrograph of deep-etched GaAs–AlGaAs quantum dots. The white dashed lines mark one of the dots. The long dashes indicate the active electron layer. (b) Transmission electron micrograph of an InAs quantum dot. The white dashed lines mark the borders of the InAs island. (With permission of S. Mendach.) (c) Schematic band structure of an InAs SAQD sample with a two-dimensional electron system (2DES) as a back contact.

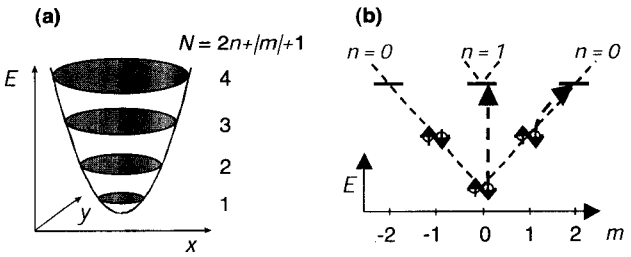


Fig. 2 (a) Schematic picture of the lateral potential of a quantum dot. (b) Energy levels of a quantum dot with parabolic potential. The curved dashed arrows indicate possible excitations of single electrons.

Therefore we will start our investigations by assuming that the corresponding single-particle energies are given by the eigenvalues of a two-dimensional harmonic oscillator

$$E_{nm} = \hbar\Omega_0(2n + |m| + 1) = N\hbar\Omega_0 \quad (1)$$

Here the two relevant quantum numbers are the radial quantum number $n = 0,1,2,\ldots$ and the azimuthal or angular-momentum quantum number $m = 0,\pm1,\pm2,\ldots$. We define a *lateral* quantum number $N = (2n + |m| + 1)$, which characterizes the $2N$-fold degenerate discrete levels, as sketched in Fig. 2a. Naturally, the presence of many other electrons in the sample leads to a screening of the potential, which acts on the test electron. Theoretically, this is described, e.g., in a self-consistent Hartree calculation of the ground state. Thus the *effective* lateral potential is flattened in the center of the dot, i.e., it becomes more and more square-well-like. This results, in a mean field approach, in a decrease of the single-particle level spacings and to a lifting of degeneracies. For simplicity, we stay for the moment in the simple picture, where we assume that also the effective, or Hartree, potential is parabolic which is a good approximation for the ground state of dots with small electron numbers. In such a single-particle picture, the electronic excitations, which can be created in the inelastic light scattering process, are transitions of electrons from occupied to unoccupied levels, as, e.g., sketched in Fig. 2b. In Fig. 2b, the energy levels for a parabolic quantum dot, which contains six electrons, is shown in more detail. The corresponding radial- and angular-momentum quantum numbers are indicated. Of course, in a real interacting system, the electronic excitations will be collective excitations, CDEs and SDEs, which are affected by Coulomb interaction.

For large electron numbers, on the order of 100 or more, it is instructive to leave the quasiatomistic picture described above for a while, and discuss the electronic excitations in terms of plasma oscillations of the electrons. For CDEs, in that case, some of the excitations are also called *confined plasmons*. For all following considerations, we assume that the *external* potential is parabolic. Consequently, the simplest CDE is a plasma oscillation where all electrons oscillate in phase in the quantum dot back and forth. This situation is, for a fixed time, schematically shown in Fig. 3a, where a dark color should

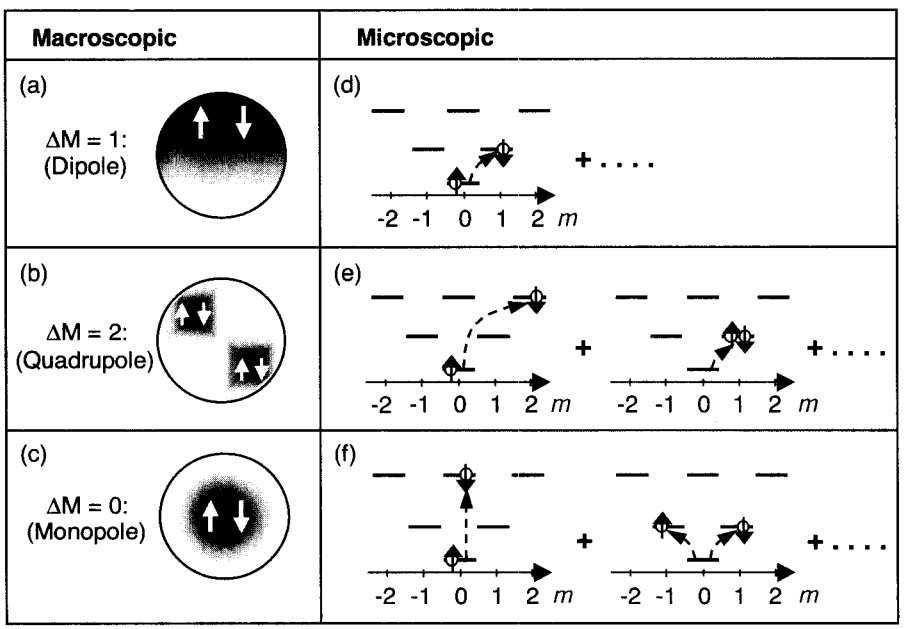


Fig. 3 Schematic drawings of the electron-density distributions for different confined plasmon modes [(a)–(c)]. Microscopic picture for the excitations of a parabolic quantum dot with two electrons [(d)–(f)].

indicate a large induced electron density. White indicates an induced positive charge distribution. Hence the induced electron density has, in this case, one node. Obviously, this excitation has a large dipole moment. It is also called the first confined plasmon or Kohn's mode, because, for the case of a parabolic external potential, its energy exactly equals the quantization energy of the external potential, and the energy of the lowest energy dipole excitation is independent of the number of electrons N in the dot. This fact is known as the so-called generalized Kohn's theorem.[33] The dipole excitation is characterized by a change of $\Delta M = 1$ of the total angular-momentum quantum number M, i.e., the angular momentum of the N-electron system. This will become clearer when we discuss the microsopic picture below. The CDEs with the next higher energies are the quadrupole excitation corresponding to $\Delta M = 2$ [Fig. 3b], and the monopole excitation with a change in total angular momentum of $\Delta M = 0$ [Fig. 3c]. The next higher excitation would again be a dipole excitation (not shown here), which is additionally accompanied by a change in the radial quantum number n. In all CDEs, the electrons with spin up and spin down oscillate in phase against the positively charged background. This is schematically indicated by the white arrows in Fig. 3. This means that the CDEs are strongly affected by the direct part of the Coulomb interaction, which in most cases results in a blueshift of the excitations. In contrast, in the corresponding SDEs, the electrons with spin up and spin down oscillate with a phase shift of π. For the *spin dipole* excitation (not shown here), corresponding to the CDE in Fig. 3a, e.g., this means that the center of mass of the charge does not oscillate. Thus the energies of the SDEs are only renormalized by exchange-correlation interaction and are therefore redshifted. We note here already that, in experiments on zincblende semiconductors, SDEs and CDEs can be distinguished by polarization selection rules:[34] CDEs are observed if the polarizations of the incoming and scattered light are parallel to each other (*polarized* geometry), and, for SDEs, the polarizations have to be perpendicular (*depolarized* geometry).

In the following, we will discuss how this macroscopic picture translates into the quasiatomistic one, which we introduced at the beginning. This makes sense for small electron numbers only, otherwise it would be too complicated. The simplest interacting system that comes to mind is a quantum dot with two electrons. In Fig. 3d–f, the dipole, quadrupole, and monopole CDEs are schematically shown for a two-electron quantum dot, respectively. More precisely, Fig. 3d–f pictorially displays the occupation of the Slater determinants, which predominantly contribute to the two-particle wavefunctions of the respective excited states. The exact many-particle wavefunction can always be written as an infinite series of single-particle Slater determinants. Fig. 3d–f exhibits the occupation of the most important Slater determinants of the respective excited states. The ground state of a two-electron dot has a total angular momentum of $M = 0$, because both electrons occupy the s level with $m = 0$. In Fig. 3, one can see that for the dipole excitation, the angular momentum of the excited state is $M = 1$, for the quadrupole excitation, $M = 2$, and, for the monopole excitation, $M = 0$. For CDEs, as displayed in Fig. 3, the spin is preserved during the excitation ($\Delta S = \Delta S_z = 0$). For the excitation of SDEs, which are triplet excitations, $\Delta S = 1$, and ΔS_z can take on values of 1, -1, and 0. This means that spinflip transitions of electrons are involved in SDEs. For zero magnetic field, $B = 0$, the three triplet excitations with $\Delta S_z = 0, \pm 1$, corresponding to the same excitation type (e.g., spin dipole, spin monopole, etc.), are degenerate. As already noted, we expect the SDEs to have lower energies than the corresponding CDEs because they experience no direct Coulomb interaction.

Experimental Details

For resonant ILS experiments, tunable lasers such as dye lasers or Ti:sapphire lasers are used. Usually, the samples are kept in a cryostat at low temperatures, typically below 10 K. A special feature of ILS spectroscopy is that a finite wave vector $\boldsymbol{q}$ can be transferred to the excitations. This is achieved by tilting the sample with respect to the incident and scattered beams. Fig. 4 schematically displays the most often used backscattering geometry, which allows the largest wave-vector transfer $\boldsymbol{q}$ into the system. Here $\boldsymbol{k}_i$ and $\boldsymbol{k}_s$ represent the wave vector of the incident and inelastically scattered photons, respectively. Under the assumption that the wavelengths of the incident, λ_i, and scattered light, λ_s, are almost equal ($\lambda_i \sim \lambda_s$, which is a good approximation because light in the visible or near-infrared range is used to excite excitations with energies in the FIR), the equation $\boldsymbol{q} \sim 4\pi\lambda_i \sin\Theta$ holds for the relation of $\boldsymbol{q}$ vs. the tilt angle Θ. The experiments discussed in this article were performed using backscattering geometry. A Ti:sapphire laser was applied for excitation, and the ILS signals were detected with a triple Raman spectrometer, equipped with a cooled, charge-coupled device detector.

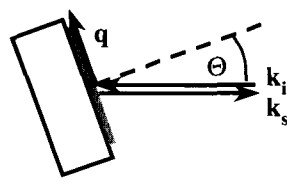


Fig. 4 Sketch of the backscattering configuration.

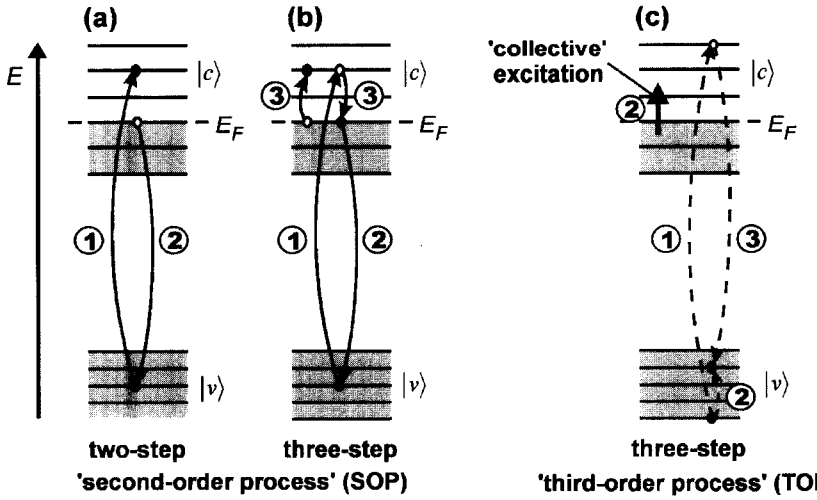


Fig. 5 Schematic pictures of the (a) two-step, (b) three-step second-order scattering processes, and (c) a three-step third-order process, which contribute to the resonant ILS in quantum dots. (From Ref. [14].)

The samples were mounted in cryostats at temperatures between $T = 2$ K and $T = 12$ K.

SCATTERING MECHANISMS

In this section, the dominant scattering mechanisms, which lead to the light scattering in low-dimensional electron systems, are qualitatively discussed. In general, the inelastic light scattering by electronic excitations in semiconductor microstructures with relatively small numbers of electrons can be observed only under specific interband resonance conditions. That means that in the scattering process, valence-band states are involved as intermediate states. Nevertheless, most of the theoretical work concerning the dynamic Raman response in microstructures is performed for nonresonant conditions, neglecting the valence-band structure. We summarize here the dominant *resonant* scattering processes, which are fragmentarily present in the literature, and which we found in our experiments on quantum wells, wires, and dots to be the dominant ones.[114]

The coupling of the radiation with the electron system is taken into account by replacing the momentum $\boldsymbol{p}$ of the electron by $\boldsymbol{p} + e\boldsymbol{A}$ in the Hamiltonian H_0 of the unperturbed system. $\boldsymbol{A}$ is the vector potential of the electromagnetic field.

$$H = \frac{1}{2m}\sum_i[(\boldsymbol{p}_i + e\boldsymbol{A}(\boldsymbol{r}_i))^2 + U(\boldsymbol{r}_i)] + V_{e-e} + V_{e-ph}$$
$$= H_0 + \frac{1}{2m}\sum_i[(\boldsymbol{A}(\boldsymbol{r}_i))^2 + \boldsymbol{p}_i\boldsymbol{A}(\boldsymbol{r}_i) + \boldsymbol{A}(\boldsymbol{r}_i)\boldsymbol{p}_i] \quad (2)$$

V_{e-e} is the Coulomb interaction, V_{e-ph} the electron–phonon interaction, and U includes the lattice-periodic potential as well as all types of external potentials. For simplicity, the spin-orbit coupling is neglected in Eq. 2. In a perturbation theory approach, the last three terms of Eq. 2 can be treated as the perturbation. Here the $\boldsymbol{A}^2$ terms give contributions to the light scattering cross section in *first* order, the $\boldsymbol{pA}$ terms in *second*-order perturbation theory. The $\boldsymbol{pA}$ terms, together with either the Coulomb interaction U_{e-e} of a photoexcited exciton with the electron system, or the electron–phonon interaction V_{e-ph} of a photoexcited exciton with the phonon bath, give contributions in *third*-order perturbation theory. V_{e-ph} leads to the scattering by phonons, which shall not be discussed here. The second-order (SOP) and third-order light scattering processes (TOP) appear to be the dominant ones in electronic Raman scattering on microstructures because they exhibit a resonant behavior.[35] In Fig. 5, transitions which contribute to the second- [Fig. 5a,b] and third-order [Fig. 5c] processes in quantum dots are sketched. Valence band levels are marked by $|v\rangle$, and conduction band levels by $|c\rangle$. In the first step of the two-step SOP, an electron is excited from a valence band level to a conduction band state. In the second step, an electron, either with the same or with opposite spin as the photoexcited electron, recombines with the hole in the valence band. Thus as a net effect, either a nonspinflip or spinflip single-particle excitation has been created in the conduction band. In the third step of the three-step SOP, another SPE is created by Coulomb interaction. This represents the screening of the SPE by other SPEs in the system. This screening also leads to the formation of collective CDEs and SDEs by direct and exchange Cou-

lomb coupling. For nonresonant conditions, the SPEs are screened by the interaction. This is formally implemented by the cancellation of the two-step SOP contribution by the three-step SOP in the scattering cross section for off-resonance conditions. In experiments on deep-etched GaAs samples with relatively large electron numbers, it was observed that under conditions of extreme resonance, which means that the laser frequency is in the vicinity of the fundamental bandgap of the underlying 2-D structure, quasi unscreened SPEs occur. They show stronger resonance enhancements than the collective excitations.[10] This means that the above-mentioned cancellation of terms does not hold in the case of extreme resonance and thus single-particle-like excitations can be observed. A quite similar behavior was reported for intraband excitations in n-type GaAs bulk samples.[36]

It was found that for laser energies well above the bandgap, an excitonic third-order scattering process, as for the case of quantum dots displayed in Fig. 5c, is the dominant one (for 2-D excitations, see Refs. [16,17]). For this scattering mechanism, which is assisted by the Coulomb interaction, in the first step the incident photon creates an exciton with a hole in the valence band and an electron in a higher conduction band level. This exciton is scattered by direct and exchange Coulomb interaction with the electron system into another state, where, e.g., as drawn in Fig. 5c, the hole is scattered into another state. By this scattering process, a collective excitation (SDE or CDE) is created in the conduction band system. In the third step, the scattered exciton recombines. This scattering process exhibits very sharp and intense resonance profiles in a plot of the scattered intensity vs. laser energy.[16] It was found that this TOP is the dominant scattering mechanism for excitation of collective SDEs and CDEs in low-dimensional GaAs samples.[10]

In the following, we elucidate these scattering processes in more detail. The scattering cross section is given by[34]

$$\frac{d^2\sigma}{d\Omega d\omega} = \frac{\omega_S}{\omega_I}\frac{e^4}{c^4 m^4} S(\omega) \tag{3}$$

where the structure factor $S(\omega)$ is defined as

$$S(\omega) = \sum_F |\langle F|V_{eff}|I\rangle|^2 \delta(E_F - E_I - \hbar\omega) \tag{4}$$

ω_I (ω_S) is the frequency of the incident (scattered) photon, and $\hbar\omega = \hbar\omega_I - \hbar\omega_S$ is the energy transfer. V_{eff} is the effective operator, which describes the transition of the system from the many-particle initial state $|I\rangle$ with energy E_I to the final state $|F\rangle$ with energy E_F. Using the second-quantization technique, this effective operator of the light scattering V_{eff} can, in a many-particle system, be expanded in terms of creation ($\hat{c}^+$) and annihilation ($\hat{c}$) operators of single-electron states[17]

$$V_{eff} = \sum_{\alpha,\beta} \gamma_{\alpha\beta} \hat{c}_\beta^+ \hat{c}_\alpha \tag{5}$$

Here each α and β represents a set of quantum numbers (n,m,σ), for the radial, azimuthal, and spin quantum numbers of the corresponding state, respectively. If we treat, for a moment, the excitons as simple electron–hole pairs without interaction, which for a quantitative analysis is certainly not correct, the scattering amplitudes $\gamma_{\alpha\beta}$ can approximately be written as

$$\gamma_{\alpha\beta} \propto \langle\alpha|e^{iqr}|\beta\rangle e_I e_S + \frac{1}{m}\sum_{\beta'} \frac{\langle\alpha|\boldsymbol{pA}_S|\beta'\rangle\langle\beta'|\boldsymbol{pA}_I|\beta\rangle}{E_\beta - E_{\beta'} + \hbar\omega_I} + \sum_{\nu,\nu'} \frac{\langle\beta|\boldsymbol{pA}_S|\nu'\rangle\langle\nu'|U_{e-e}|\nu\rangle\langle\nu|\boldsymbol{pA}_I|\alpha\rangle}{(E_\beta - E_{\nu'} + \hbar\omega_S)(E_\alpha - E_\nu - \hbar\omega_I)} \tag{6}$$

where, in the last two terms, we have written only the strongest resonant term. The first term in Eq. 6 represents the nonresonant contributions from the $\boldsymbol{A}^2$ terms in first-order perturbation theory. Here the scalar product of the polarization vectors of the incident ($\boldsymbol{e}_I$) and scattered ($\boldsymbol{e}_S$) photon shows that this nonresonant term yields only scattering by plasmons (CDEs). Spinflip processes, which are necessary for the excitation of SDEs. are only possible in the resonant second and third terms. The second term describes second-order processes, as schematically shown in Fig. 5a, and the third term represents third-order resonant scattering, as displayed in Fig. 5c.

To illustrate this, Fig. 6 shows experimental ILS spectra of electronic excitations in deep-etched GaAs quantum dots with 240 nm geometrical diameter and about 200 electrons per quantum dot. The spectra in the lower part were recorded at laser energies well above the effective bandgap. In the lowest spectrum at a laser energy of E_L = 1587 meV, three peaks can be observed, which can be identified as SDEs because of polarization selection rules. (We will come to the detailed interpretation of the excitations in the section "Experiments on GaAs–AlGaAs Deep-Etched Quantum Dots.") Correspondingly, the peaks in the polarized spectrum can be identified as CDEs. We have found that CDEs, in particular, exhibit a very sharp [full width at half-maximum (FWHM)$\sim$3.5 meV] resonance profile in this range of laser energies—which leads us to the interpretation that this scattering by *collective* excitations is caused by the third-order TOP, as discussed before. If the laser frequency is slightly lowered (spectra at E_L = 1585 meV in Fig. 6), additional broad features appear, especially in the polarized spectrum, which then slowly evolve into intense peaks, by further lowering E_L toward the bandgap. These features are marked in Fig. 6 by dotted arrows. Under

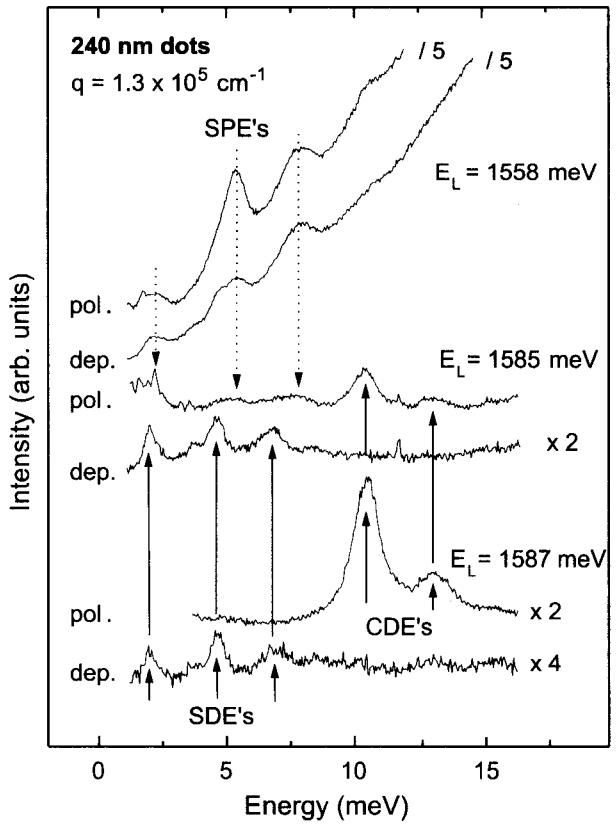


Fig. 6 Polarized and depolarized ILS spectra of deep-etched GaAs–AlGaAs quantum dots for different laser energies E_L. (From Ref. [37].)

conditions of extreme resonance (e.g., E_L=1558 meV in Fig. 6), these peaks dominate the spectra and are present in both polarization configurations. Therefore we interpret them, within the framework described above, as SPEs, which are caused by an SOP under conditions of extreme resonance. We note that the observation of similar SPEs in multilayered quantum dots were reported by Lockwood et al.[9] From the experimental spectra in Fig. 6, we can directly deduce for the corresponding collective SDEs and CDEs the energy renormalizations due to many-particle interactions.[10]

EXPERIMENTS ON GaAs–AlGaAs DEEP-ETCHED QUANTUM DOTS

In this section, we focus on ILS experiments on modulation-doped GaAs dots in more detail. In particular, we will discuss parity selection rules for the collective excitations and the importance of resonant excitation.

Fig. 7 shows experimental ILS spectra of quantum dots with approximately 200 electrons per dot.[37] In these experiments in backscattering geometry, the wave-vector transfer q parallel to the plane of the dot was close to zero. In each scattering configuration, depolarized and polarized, there is one mode visible. On general grounds, one can say that in a symmetric system, the allowed modes have even parity because the ILS process is a two-photon process. This is in contrast to direct absorption, which is a one-photon process. Consequently, the observed SDE and CDE are even-parity modes. Referring to the section "Characteristics of Quantum Dots and Experimental Details," we find that the lowest energy modes with even parity are the monopole modes with $\Delta M = 0$. The modes are labeled $(\Delta n,\Delta m)$ corresponding to the changes in radial (Δn) and angular-momentum (Δm) quantum numbers, respectively, of the involved transitions. The assignment of the SDE was deduced from its magnetic field behavior[11] (not shown here) and the CDE is considered in more detail below. We want to note here that the polarized spectrum demonstrates the exclusion principle between FIR and Raman spectroscopy: the FIR-allowed mode $CDE_{(0,1)}$, the Kohn's mode, has odd parity and hence a large dipole moment (in the section "Characteristics of Quantum Dots and Experimental Details"). It has, for the sample displayed in Fig. 7, an energy of about 6 meV and is not visible at all in the Raman spectrum (Fig. 7). On the other hand, the Raman-allowed monopole mode [$CDE_{(1,0)}$] has no dipole moment at all. Considering that the induced density of this mode has only nodes in radial direction [see Fig. 3c], it can also be regarded as a so-called breathing mode. Thus the experiments demonstrate that, for resonant ILS on

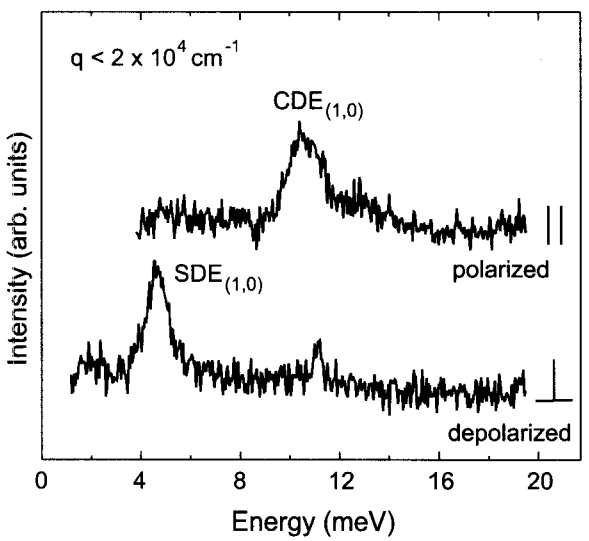


Fig. 7 Polarized and depolarized ILS spectra of 240 nm GaAs–AlGaAs quantum dots with approximately 200 electrons per dot. (From Ref. [37].)

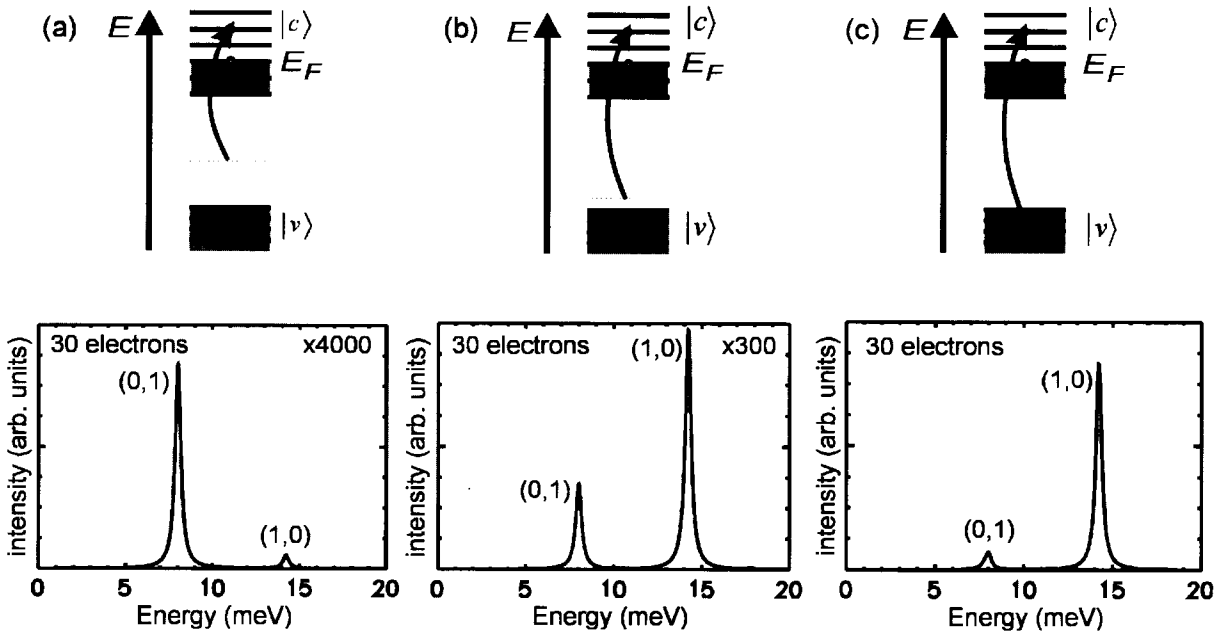


Fig. 8 Calculated ILS spectra for a quantum dot with 30 electrons for (a) off-resonance, (b) close-to-resonance, and (c) resonant conditions. The schematic pictures show the different resonance conditions. The curved arrows indicate the laser energies. (From Ref. [37].)

circularly symmetric dots at approximately zero wave-vector transfer, the general parity selection rules hold.

In the following, we will elucidate these parity selection rules from the theoretical point of view in more detail. Most of the calculations of ILS spectra for low-dimensional electron systems have been carried out for nonresonant conditions, because in such cases, one does not have to deal with the complex valence-band states. Fig. 8 shows a series of calculated resonant ILS spectra for off-resonance and for different resonance conditions.[37] The calculations have been performed for a 30-electron quantum dot with $\hbar\Omega_0 = 8$ meV for the external potential. In Fig. 8a, the laser energy was chosen to be far away from resonance, i.e., off-resonance conditions. Under experimental conditions, in this regime the scattered intensities are much too low to be observable. Note that the calculated spectrum in Fig. 8a is multiplied by a factor of 4000 to be of comparable strength with the resonant spectrum in Fig. 8c. One can observe in Fig. 8a that, in the off-resonance case, the parity selection rules have completely changed: for the sake of curiosity, the FIR-allowed Kohn's mode [$CDE_{(0,1)}$] is the dominant mode. As the laser energy approaches resonance, the situation changes, and, under resonance conditions [Fig. 8c], the experimentally observed parity selection rules are confirmed, i.e., the monopole mode $CDE_{(1,0)}$ is the dominant mode in the polarized spectrum. This clearly demonstrates that, for a correct description of the relative intensities of the excitations, the resonant scattering process is crucial.

An at least partial breakdown of the parity selection rules can be achieved by the transfer of a finite wave vector $\boldsymbol{q}$ parallel to the plane of the dot. Fig. 9 shows a series of depolarized and polarized spectra for the same dots in Fig. 7 for different wave-vector transfer $\boldsymbol{q}$. As $\boldsymbol{q}$ increases, symmetry-forbidden modes (the spin dipole mode $SDE_{(0,1)}$, and a higher dipole mode $CDE_{(1,1)}$) gain relative intensity. Furthermore, the spin quadrupole mode $SDE_{(0,2)}$—which, for a parabolic effective potential, would be energetically degenerate with the spin monopole mode $SDE_{(1,0)}$ if interaction effects are neglected—becomes visible.[11] A so-far unsolved puzzle focuses on why the violation of parity selection rules with wave-vector transfer $\boldsymbol{q}$ is much stronger for SDEs than for CDEs.

EXPERIMENTS ON InAs SELF-ASSEMBLED QUANTUM DOTS

Few-Electron Quantum-Dot Atoms

It has been demonstrated that, by application of metallic gates, InAs SAQD can be charged with single electrons. For experiments that will be described below, a special sample design, where a 2-D electron system was used as a

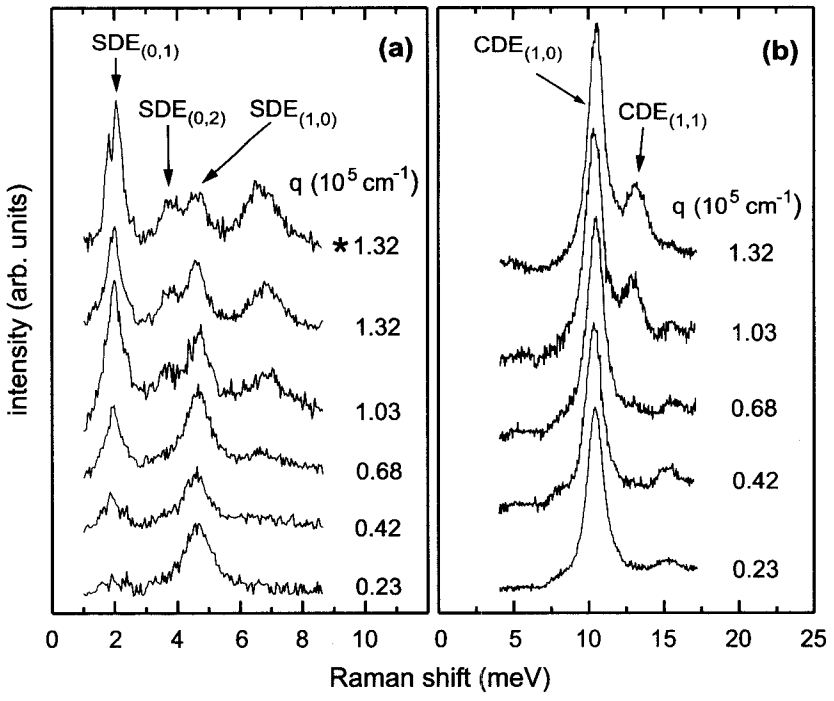


Fig. 9 (a) Depolarized and (b) polarized ILS spectra of GaAs quantum dots for different wave-vector transfer q. (From Ref. [11].)

back contact, was applied (for details, see Ref. [30]). The charging of dots can be monitored in situ by capacitance measurements (Fig. 10, inset). For the ILS experiments, resonant excitation was achieved by tuning the laser to the $E_0 + \Delta$ gap of the InAs SAQD ($\sim$1.65 eV), which is far above the fundamental PL transition energies of the structure ($\sim$1.1–1.2 eV). Fig. 10 displays a series of polarized ILS spectra. At E = 33.4 meV and E = 36.6 meV, two sharp lines can be observed that result from the TO- and LO-phonon excitations of the GaAs bulk material in the structure. At higher energies, broader bands are visible in the range 45–55 meV (labeled as A and B), which we attribute to CDEs of the electrons in the quantum dots. Both the positions and the linewidths of these bands change with applied gate voltage V_{Gate}. The capacitance trace, displayed in the inset of Fig. 10, shows that, by varying V_{Gate}, the s- and the p-shell of the quantum dots can be charged with electrons. The doublet structure around $V_{\text{Gate}} = -0.05$ V stems from the charging of the s shell with two electrons, and, at the broad plateau at positive V_{Gate}, the p-shell is loaded with four electrons. The bands A in Fig. 10 were interpreted to be a result of transition of electrons from the s to the p shell (s–p transitions) of the quasiatoms, and the B bands from p–d transitions:[30] it was assumed that 2–3 confined single-particle energy levels exist in the quantum dots, which is schematically shown in the drawings on the right-hand side of Fig. 10. Here three selected situations with $N = 2$, $N = 4$, and $N = 6$ electrons per quantum dot are sketched in a single-particle picture. The vertical arrows indicate possible transitions of single electrons. The SAQD stick out of the InAs wetting layer. Therefore the dot potentials are flattened at the edges because of the wetting-layer continuum. This leads to the situation that the single-particle transition, sketched as B in the inset of Fig. 10, has a smaller energy than the transition A. So far, in the interpretation, only transitions between single-particle states have been considered. Within this simplified picture, one is not able to explain the observed shift of excitation A (s–p transitions) with increasing electron number N in the quantum dots, as observed in Fig. 10. In the voltage range where the p-shell is filled with electrons ($V_{\text{Gate}} \sim$ 0.2–0.6 V), we find, in Fig. 10, a shift of A to lower energies and a broadening of the transition.

More accurately, the experimentally observed excitations are of course collective CDEs of the few-electron quantum dots, and are therefore affected by Coulomb interaction. To take account of this, model calculations where the excitation energies were calculated by using exact numerical diagonalization of the Hamilton operator

$$H = \sum_{i=1}^{N}\left[\frac{\boldsymbol{p}_i^2}{2m^*} + \frac{m^*}{2}\Omega_0^2 \boldsymbol{r}_i^2\right] + \frac{e^2}{4\pi\varepsilon\varepsilon_0}\sum_{i\neq j}^{N}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \qquad (7)$$

of the N-electron quantum dot were performed.[30] There, a two-dimensional quantum dot with a parabolic confining

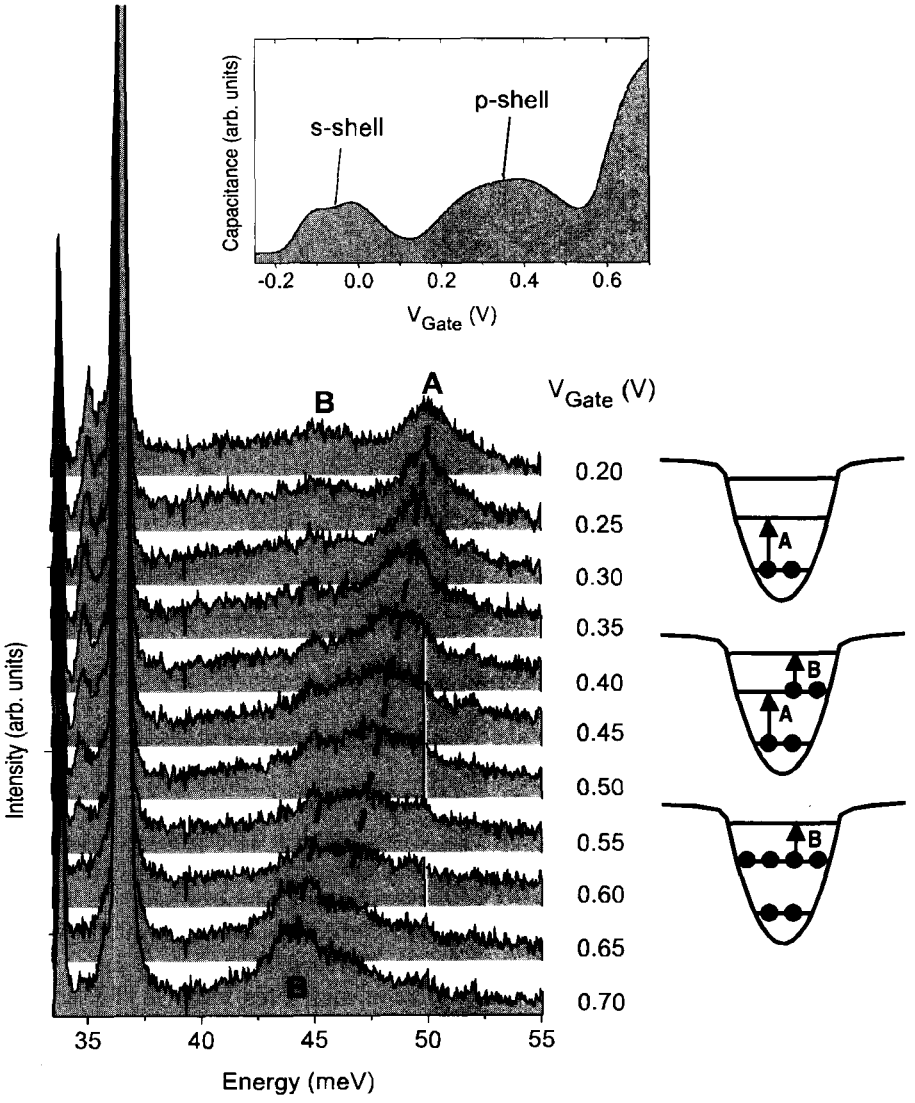


Fig. 10 Series of polarized ILS spectra of InAs SAQD for different gate voltages V_{Gate}. The upper inset shows a capacitance measurement of the sample. On the right-hand side, different electron configurations in the dots are schematically shown. (From Ref. [30].)

potential in lateral direction [$\boldsymbol{r} = (r_x, r_y)$] with quantization energy $\hbar\Omega_0 = 50$ meV was assumed. By exact numerical diagonalization, the low-energy excitations of the N-electron quantum dot were determined, where, during the excitation, the total spin is preserved. This selects the excitations, which one expects in polarized ILS spectra (CDEs). Fig. 11 exhibits the calculated excitation energies for electron numbers $N = 2$–6 in the quantum dot. The vertical bars mark the energetic positions of the few-particle excitations. First, one can see that, independent of N, there is always an excitation at the energy $\hbar\Omega_0 = 50$ meV of the external confining potential. This represents the Kohn's mode. However, for $N > 2$, in the calculations additional mode energies appear below the energy of the Kohn's mode. These are a consequence of the different possible final states of the electrons after excitation. For very small electron numbers, these different final states make a significant difference in the energy of the state. For $N = 3$, e.g., the configuration where two electrons in the p shell have antiparallel spin has the lowest energy. With increasing electron numbers, more complex electron configurations occur that can have different energies. The most important result of Fig. 11, concerning the experiments, is that with increasing N, the spectral weight of the low-energy excitations shifts to lower energies. This might ascribe the shift and broadening of the s–p transitions, as observed in the experiments (band A in Fig. 10), to be due to additional excitations at lower energies, which cannot be individually resolved in the ensemble experiment. Clearly, the model can only

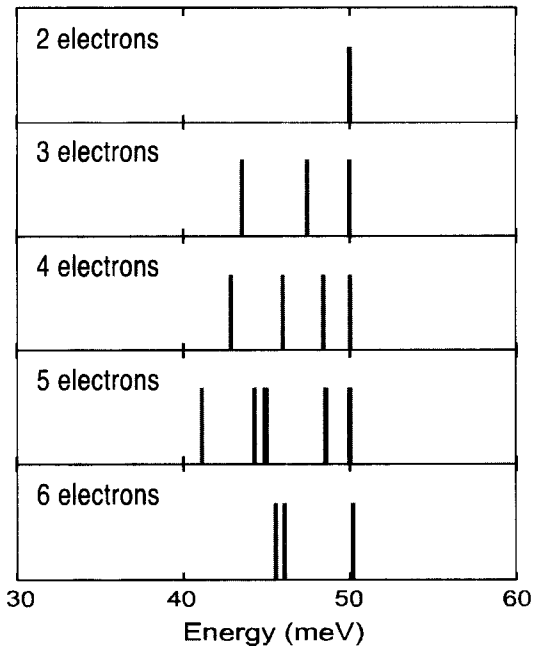


Fig. 11 Calculated few-electron excitation energies for a parabolic InAs quantum dot. (From Ref. [30].) (*View this art in color at www.dekker.com.*)

qualitatively explain the experiments. The most important limitation is the assumption of a parabolic potential, which deviates from the more realistic form (sketched in Fig. 10) and which cannot account for the lower energy of band B. However, one should note that the important qualitative result, i.e., the appearance of additional collective modes at lower energies for increasing electron number, is independent of the choice of parameters and the form of the confining potential. This leads to the conclusion that these results might generally apply for quantum-dot atoms.

CONCLUSION

There is a variety of theoretical and experimental literature on ILS spectroscopy of electronic excitations in quantum dots. Most of the experiments have been performed on modulation-doped GaAs–AlGaAs quantum dots. In these mesoscopic systems with electron numbers on the order of 100 per quantum dot, CDEs, SDEs, and SPEs could be observed. More sophisticated theoretical models demonstrated, in comparison with experiments, the importance of the resonant scattering process, including valence-band states as intermediate states, for a correct description of, e.g., parity selection rules. Recently, resonant ILS experiments on InAs SAQD with tunable electron numbers have been reported. This opens the road for the investigation of excitations in few-electron quantum dot atoms. With the prospect of single-dot Raman spectroscopy, in the future one might be able to observe and resolve electronic excitations as individual sharp lines, caused by interaction effects of few electrons in artificial atoms.

ACKNOWLEDGMENTS

I would like to thank many colleagues who contributed to our work, as listed in the references, and B. Wunsch for providing the theoretical results, which were used to prepare Fig. 3d–f. In particular, I am grateful to D. Heitmann for his support and critical reading of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft via SFB 508 and a Heisenberg Grant (SCHU 1171/2).

REFERENCES

1. Reed, M.A.; Randall, J.N.; Aggarwal, R.J.; Matyi, R.J.; Moore, T.M.; Wetsel, A.E. Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure. Phys. Rev. Lett. **1988**, *60* (6), 535–537.
2. Hansen, W.; Smith, T.P., III; Lee, K.Y.; Brum, J.A.; Knoedler, C.M.; Hong, J.M.; Kern, D.P. Zeeman bifurcation of quantum-dot spectra. Phys. Rev. Lett. **1989**, *62* (18), 2168–2171.
3. Sikorski, C.; Merkt, U. Spectroscopy of electronic states in InSb quantum dots. Phys. Rev. Lett. **1989**, *62* (18), 2164–2167.
4. Demel, T.; Heitmann, D.; Grambow, P.; Ploog, K. Nonlocal dynamic response and level crossings in quantum-dot structures. Phys. Rev. Lett. **1990**, *64* (7), 788–791.
5. Lorke, A.; Kotthaus, J.P. Coupling of quantum dots on GaAs. Phys. Rev. Lett. **1990**, *64* (21), 2559–2562.
6. Meurer, B.; Heitmann, D.; Ploog, K. Single-electron charging of quantum-dot atoms. Phys. Rev. Lett. **1992**, *68* (9), 1371–1374.
7. Bollweg, K.; Kurth, T.; Heitmann, D.; Gudmundsson, V.; Vasiliadou, E.; Grambow, P.; Eberl, K. Detection of compressible and incompressible states in quantum dots and antidots by far-infrared spectroscopy. Phys. Rev. Lett. **1996**, *76* (15), 2774–2777.
8. Strenz, R.; Bockelmann, U.; Hirler, F.; Abstreiter, G.; Böhm, G.; Weimann, G. Single-particle excitations in quasi-zero- and quasi-one-dimensional electron systems. Phys. Rev. Lett. **1994**, *73* (22), 3022–3025.
9. Lockwood, D.J.; Hawrylak, P.; Wang, P.D.; Sotomayor Torres, C.M.; Pinczuk, A.; Dennis, B.S. Shell

structure and electronic excitations of quantum dots in a magnetic field probed by inelastic light scattering. Phys. Rev. Lett. **1996**, *77* (2), 354–357.
10. Schüller, C.; Biese, G.; Keller, K.; Steinebach, C.; Heitmann, D.; Grambow, P.; Eberl, K. Single-particle excitations and many-particle interactions in quantum wires and dots. Phys. Rev., B **1996**, *54* (24), R17304–R17307.
11. Schüller, C.; Keller, K.; Biese, G.; Ulrichs, E.; Rolf, L.; Steinebach, C.; Heitmann, D.; Eberl, K. Quasi-atomic fine structure and selection rules in quantum dots. Phys. Rev. Lett. **1998**, *80* (12), 2673–2676.
12. Biese, G.; Schüller, C.; Keller, K.; Steinebach, C.; Heitmann, D.; Grambow, P.; Eberl, K. Coupling of lateral and vertical electron motion in GaAs–AlGaAs quantum wires and dots. Phys. Rev., B **1996**, *53* (15), 9565–9567.
13. Pinczuk, A.; Schmitt-Rink, S.; Danan, G.; Valladares, J.P.; Pfeiffer, L.N.; West, K.W. Large exchange interactions in the electron gas of GaAs quantum wells. Phys. Rev. **1989**, *63* (15), 1633–1636.
14. Schüller, C. Raman Spectroscopy of Quantum Dots. In *Festkörperprobleme/Advances in Solid State Physics*; Kramer, B., Ed.; Vieweg: Braunschweig/Wiesbaden, 1999; Vol. 38, 167–181.
15. Das Sarma, S.; Wang, D.-W. Resonant Raman scattering by elementary electronic excitations in semiconductor structures. Phys. Rev. Lett. **1999**, *83* (4), 816–819.
16. Danan, G.; Pinczuk, A.; Valladares, J.P.; Pfeiffer, L.N.; West, K.W.; Tu, C.W. Coupling of excitons with free electrons in light scattering from GaAs quantum wells. Phys. Rev., B **1989**, *39* (8), 5512–5515.
17. Govorov, A.O. Resonant light scattering induced by Coulomb interaction in semiconductor microstructures. J. Phys. Condens. Matter **1997**, *9*, 4681–4690.
18. Steffens, O.; Suhrke, M. Novel spin features in Raman spectra of few-electron quantum dots. Phys. Rev. Lett. **1999**, *82* (19), 3891–3894.
19. Pi, M.; Barranco, M.; Emperador, A.; Lipparini, E.; Serra, Ll. Current–density-functional approach to large quantum dots in intense magnetic fields. Phys. Rev., B **1998**, *57* (23), 14783–14792.
20. Serra, Ll.; Barranco, M.; Emperador, A.; Pi, M.; Lipparini, E. Spin and density longitudinal response of quantum dots in the time-dependent local-spin-density approximation. Phys. Rev., B **1999**, *59* (23), 15290–15300.
21. Lipparini, E.; Barranco, M.; Emperador, A.; Pi, M.; Serra, Ll. Transverse dipole spin modes in quantum dots. Phys. Rev., B **1999**, *60* (12), 8734–8742.
22. Mooradian, A. Light scattering from single-particle electron excitations in semiconductors. Phys. Rev. Lett. **1968**, *20* (20), 1102–1104.
23. Blum, F.A. Inelastic light scattering from semiconductor plasmas in a magnetic field. Phys. Rev., B **1970**, *1* (3), 1125–1135.
24. Sassetti, M.; Kramer, B. Resonant Raman scattering by collective modes of the one-dimensional electron gas. Phys. Rev. Lett. **1998**, *80* (7), 1485–1488.
25. Wang, D.W.; Millis, A.J.; Das Sarma, S. Where is the Luttinger liquid in one-dimensional semiconductor quantum wire structures? Phys. Rev. Lett. **2000**, *85* (21), 4570–4573.
26. Steinebach, C.; Schüller, C.; Heitmann, D. Single-particle-like states in few-electron quantum dots. Phys. Rev., B **2000**, *61* (23), 15600–15602.
27. Gammon, D.; Steel, D.G. Optical studies of single quantum dots. Phys. Today October, **2002**, *55* (10), 36–41 and references therein.
28. Drexler, H.; Leonard, D.; Hansen, W.; Kotthaus, J.P.; Petroff, P.M. Spectroscopy of quantum levels in charge-tunable InGaAs quantum dots. Phys. Rev. Lett. **1994**, *73* (16), 2252–2255.
29. Chu, L.; Zrenner, A.; Bichler, M.; Böhm, G.; Abstreiter, G. Raman spectroscopy of In(Ga)As/GaAs quantum dots. Appl. Phys. Lett. **2000**, *77* (24), 3944–3946.
30. Brocke, T.; Bootsmann, M.-T.; Wunsch, B.; Tews, M.; Pfannkuche, D.; Heyn, Ch.; Hansen, W.; Heitmann, D.; Schüller, C. Inelastic light scattering on few-electron quantum-dot atoms. Physica E. *in press*.
31. Wojs, A.; Hawrylak, P.; Fafard, S.; Jacak, L. Electronic structure and magneto-optics of self-assembled quantum dots. Phys. Rev., B **1996**, *54* (8), 5604–5608.
32. Merkt, U. Far-infrared spectroscopy of quantum dots. Physica B **1993**, *189*, 165–175.
33. Maksym, P.A.; Chakraborty, T. Quantum dots in a magnetic field: Role of electron–electron interactions. Phys. Rev. Lett. **1990**, *65* (1), 108–111.
34. Hamilton, D.; McWhorter, A.L. Raman Scattering from Spin-Density Fluctuations in n-GaAs. In *Light Scattering Spectra of Solids*; Wright, G.B., Ed.; Springer: New York, 1969; 309.
35. Burstein, E.; Pinczuk, A.; Mills, D.L. Inelastic light scattering by charge carrier excitations in two-dimensional plasmas: Theoretical considerations. Surf. Sci. **1980**, *98*, 451–468.
36. Pinczuk, A.; Brillson, L.; Burstein, E. Resonant light scattering by single-particle electronic excitations in n-GaAs. Phys. Rev. Lett. **1971**, *27* (6), 317–320.
37. Schüller, C.; Steinebach, C.; Heitmann, D. Spin- and charge-density excitations in quantum dots: A Raman study. Solid State Commun. **2001**, *119*, 323–331.

Quantum Dots Made of Cadmium Selenide (CdSe): Formation and Characterization

Kenzo Maehashi
Hisao Nakashima
Osaka University, Osaka, Japan

INTRODUCTION

II–VI semiconductors such as ZnSe, ZnS, and CdSe have wider band gap and larger exciton binding energies as compared with III–V semiconductors. Excitons of II–VI semiconductors still exist at room temperature. Therefore II–VI semiconductor devices indicate a possibility of new devices using excitons, such as optical modulators and self-electro-optic effect devices.[1–5] Furthermore, in the field of optical data transmission, green laser is suitable for short- and medium-range communication purposes using plastic optical fibers with polymethyl methacrylate cores, which have an advantage of lower costs than silica fibers. Therefore II–VI semiconductors are also optimum candidates for light source of optical data transmission through plastic optical fibers.

Semiconductor quantum structures have been intensively investigated since Esaki and Tsu[6] proposed novel artificial superlattice structures. Recently, quantum dots (QDs) have attracted much attention for the optoelectronic device applications and fundamental physics because they provide zero-dimensional structures with δ-function density of states, which dramatically improve performances of optoelectronic devices such as semiconductor lasers. To fabricate lower-dimensional structures, great efforts have been made using various methods, such as selective epitaxial growth, lithography, etching, etc.[7–12] However, these techniques give film damages, such as defects and contamination. On the other hand, self-organized QDs have an advantage of fabrication of high-density and high-quality QDs. Especially, InAs QDs of III–V materials have been known to be formed on GaAs surfaces, which are followed by the two-dimensional growth of InAs wetting layer, because InAs has a larger lattice constant by 7% than GaAs.[13–18]

CdSe/ZnSe system is expected to naturally form QDs because of large lattice mismatch of about 7% between ZnSe and CdSe. In this session, we have investigated the formation and optical properties of self-organized CdSe QDs on ZnSe (001) surfaces with the use of photoluminescence (PL) and transmission electron microscopy (TEM) measurements. Moreover, applying the QD system to optoelectronical devices to understand carrier dynamics and energy structures of QDs is very important. For example, relaxation mechanism of InAs QDs has been reported by PL excitation measurements.[19,20] As compared with InAs QDs, CdSe QDs have stronger electron–phonon interactions and much larger band-gap energy. Then, the carrier relaxation mechanisms in self-organized CdSe QDs are more interesting subjects. In this study, we have also investigated optical properties of self-organized CdSe QDs by selectively excited PL measurements.

GROWTH OF CdSe/ZnSe STRUCTURES

ZnSe/CdSe/ZnSe structures were fabricated by molecular beam epitaxy (MBE) on GaAs (001) substrates after GaAs buffer layer growth. Compound sources of ZnSe and CdSe, whose purities are 6N, were used for ZnSe and CdSe growth, respectively.[21] In the case of compound materials, the temperatures necessary to obtain beam enough flux intensity for MBE growth are much higher as compared with the temperatures needed in the conventional case for columns II and VI elemental sources. As a result, it is convenient to control the beam flux and to precisely control the composition beam flux ratio. In addition, because only molecules of Se_2 are generated, the sticking coefficient of molecules is enhanced as compared with elemental sources. Therefore high-quality films are expected to be obtained.

The polished GaAs (001) substrates were etched in a 4:1:1 solution of H_2SO_4:H_2O:H_2O_2 before placing them in the MBE loading chamber with In backing on Mo blocks. The substrates were annealed at 580°C with As irradiation to remove native oxide. First, the 200-nm-thick GaAs buffer layer was grown at 580°C in the MBE system. Next, ZnSe layer (100 nm) was grown on the GaAs surfaces at 290°C. When ZnSe growth starts on the GaAs surfaces, the GaAs (2 × 4) reconstruction pattern disappeared immediately and the reflection high-energy electron diffraction (RHEED) intensity oscillation starts, as shown in Fig. 1. The RHEED intensity oscillation during ZnSe growth on GaAs surfaces continued until more than 100-nm thickness of ZnSe film. This indicates

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013669

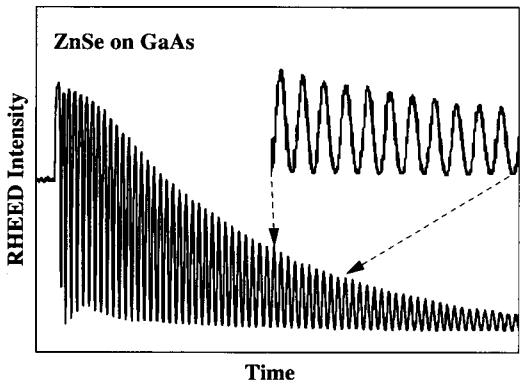


Fig. 1 RHEED intensity oscillation of ZnSe growth on GaAs (001) surfaces.

that ZnSe proceeds in layer-by-layer mood on GaAs surfaces and that flat interfaces are formed between ZnSe films and GaAs surfaces. The RHEED patterns during ZnSe growth showed both (2 × 1) and c(2 × 2) reconstruction patterns. Therefore ZnSe films with high quality were obtained.[22]

Then, a few monolayers (MLs) of CdSe were grown on the ZnSe surfaces at 290°C. The growth rate of CdSe was 1.1 ML/min, which was measured by RHEED intensity oscillation. To confirm the formation of self-organized CdSe QDs, the bulk spot intensity of RHEED was monitored during the growth of CdSe with the use of both a combination of a charge-coupled device (CCD) camera and a photo diode.[23] Finally, ZnSe layer was grown after the CdSe growth without growth interruption.

FORMATION OF CdSe QUANTUM DOTS

Fig. 2 shows PL spectra measured for the samples with various CdSe coverages.[24] Photoluminescence measurement was carried out at liquid He temperature using a 32-cm grating monochromator with a cooled CCD camera. A He–Cd laser line of 325 nm was used as an excitation light. With increasing the CdSe thickness from 1.3-ML coverages, the PL peaks shift to lower energy, which is a result of the quantum size effect. The small jump in peak position between 2.2 and 2.4 ML CdSe coverages is also observed. Photoluminescence intensity and full widths at half maximum (FWHMs) gradually increase with increasing the CdSe thickness from 1.3-ML coverages. The maximum PL intensity is obtained at about 2.2 ML CdSe coverages. However, at more than 2.4 ML CdSe coverages, PL intensities dramatically decrease, as shown in Fig. 2. The emission is hardly observed at 2.6 and 3.1 ML CdSe coverages.

Next, cross-sectional TEM images are observed for the samples with various CdSe coverages. The cross-sectional TEM specimens were prepared by mechanical polish and dimpling followed by ion etching. The TEM images were observed by JEOL JEM-3000F electron microscope at 300 kV. Fig. 3a shows a cross-sectional TEM image of CdSe 1.6-ML coverages taken along the [011] direction. The dark region shows the CdSe layer. A lot of CdSe QDs are clearly observed, indicating that CdSe QDs are naturally formed on ZnSe surfaces. Any defects are not found in this image. Fig. 3b shows a cross-sectional high-resolution TEM image of single CdSe QD. Single CdSe QD with several nanometers size is observed. Kirmse et al.[25] have reported with plan-view TEM images that two classes of CdSe QDs coexist with an average lateral size of < 10 nm (area density of 100 μm^{-2}) and 10–50 nm (20 μm^{-2}), respectively. However, we did not observe such large CdSe QDs. As compared with the atomic force microscopy image in Ref. [26], the diameter of CdSe QDs in Fig. 3 is smaller and the density is much higher. Fig. 4a shows a cross-sectional TEM image of 2.6-ML coverages taken along the $[0\bar{1}1]$ direction. This reveals that much stacking faults are observed in the ZnSe capping layer. It is found that these stacking faults have an origin at around CdSe / ZnSe interfaces. Fig. 4b shows a cross-sectional high-resolution TEM image at the interface between CdSe / ZnSe. The stacking faults start from the interface between ZnSe capping layer and CdSe QDs.

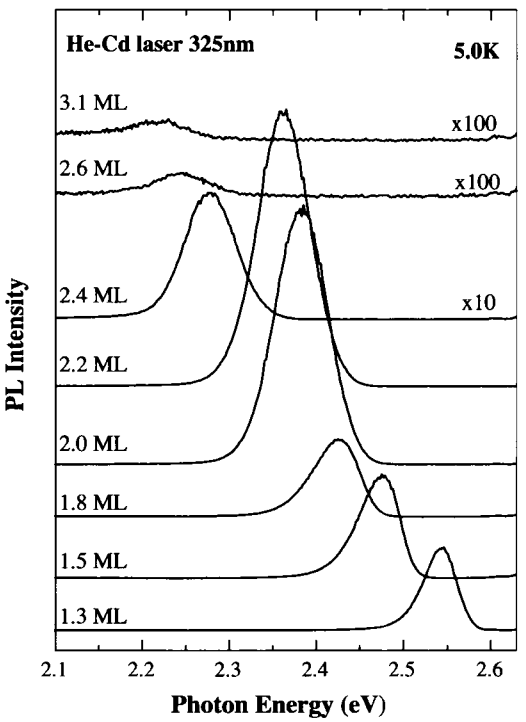


Fig. 2 Photoluminescence spectra as a function of CdSe thickness. (From Ref. [24].)

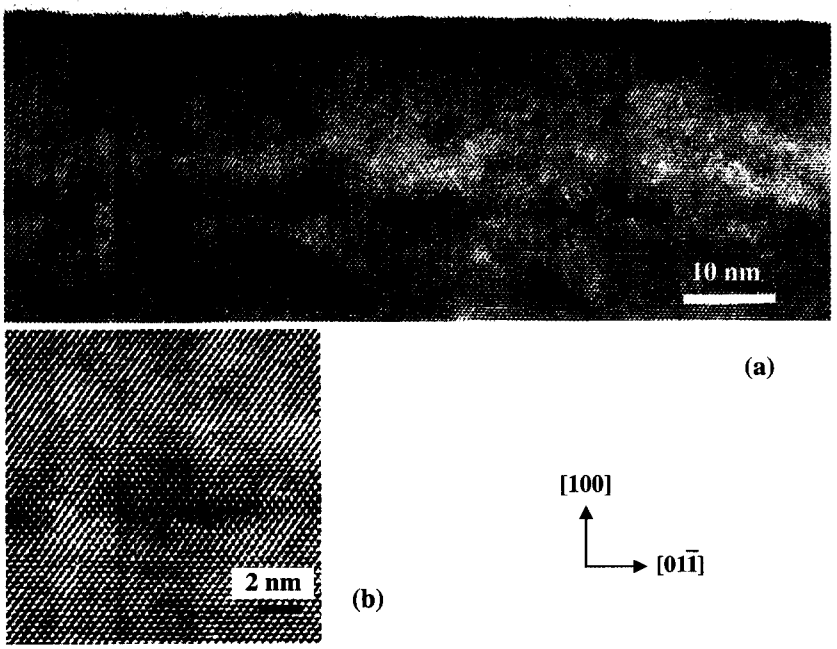


Fig. 3 (a) Cross-sectional TEM image of CdSe 1.6-ML coverage taken along the [011] direction and (b) cross-sectional high-resolution TEM images of single CdSe QD.

Therefore stacking faults are considered to be induced by large stress after ZnSe cap layer growth.

Photoluminescence and TEM results indicate that the PL intensity at less than 2.2 ML CdSe coverages shown in Fig. 2 comes from CdSe QDs. This is consistent with the result that the defects in the samples reduce dramatically the luminescence intensity of thick samples, as shown in Figs. 2 and 4. Furthermore, CdSe QDs showed a strong green emission even at room temperature.[27] These results indicate that CdSe QDs with high quality are formed on ZnSe surfaces.

To investigate the optical properties for the effect of the QD shape, polarization dependence of PL from CdSe QDs is measured.[28] Fig. 5 shows PL spectra of 2.0 ML CdSe coverages at 5 K. The solid curve and the dashed curve are spectra of polarization components, which are

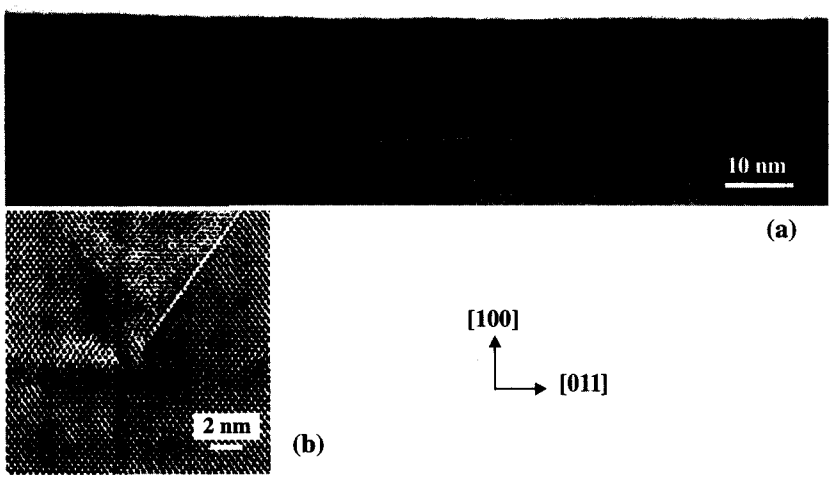


Fig. 4 (a) Cross-sectional TEM image of CdSe 2.6-ML coverage taken along the [0$\bar{1}$1] direction and (b) cross-sectional high-resolution TEM image at the interface between CdSe / ZnSe.

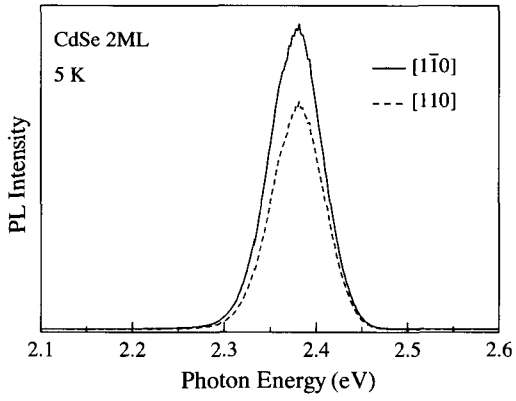


Fig. 5 Photoluminescence polarization dependence of CdSe QDs at 5 K. (From Ref. [28].)

parallel to the [1$\bar{1}$0] direction and to the [110] direction, respectively, as shown in Fig. 5. A peak at 2.38 eV comes from the CdSe QDs, and no peak from wetting layers is observed. The intensity for the polarization parallel to the [1$\bar{1}$0] direction ($I[1\bar{1}0]$) is larger than that to the [110] direction ($I[110]$), as shown in Fig. 5. No polarization anisotropy was observed between the [100] and [010] directions. Therefore the PL spectra are strongly polarized parallel to the [1$\bar{1}$0] direction. The PL peaks shown in Fig. 5 come from the heavy-hole transition, and the strong polarization is considered to be a result of the heavy-hole confinement in CdSe QDs, indicating that lateral confinement along the [1$\bar{1}$0] direction is weaker than that along the [110] direction. Consequently, CdSe QDs are considered to have the longer axis along the [1$\bar{1}$0] direction than the [110] direction.

From Fig. 5, a PL polarization degree of CdSe QDs was estimated to be 14.6%, which is defined by $(I[1\bar{1}0]-I[110])/(I[1\bar{1}0]+I[110])$. We also discuss about the CdSe coverage dependence of polarization PL spectra. Fig. 6 shows the PL polarization degrees as a function of CdSe coverage from 1.5 to 2.4 ML at 5 K. The PL spectrum for each CdSe coverage was clearly polarized parallel to the [1$\bar{1}$0] direction, indicating that the shape of QDs for each CdSe coverage has the long axis along the [1$\bar{1}$0] direction. The PL polarization degrees increase with increasing CdSe coverage, as shown in Fig. 6. The result reveals that the PL polarization degrees increase with QD formation, suggesting that the ratio of length along the [1$\bar{1}$0] direction to that along the [110] direction increases with CdSe coverage. At 2.2-ML coverages, PL spectrum has maximum polarization degree. At 2.4 ML CdSe coverages, defects were induced and, simultaneously, polarization degree decreases as shown in Fig. 6, which is considered to be a result of the release from the stress in CdSe QDs.

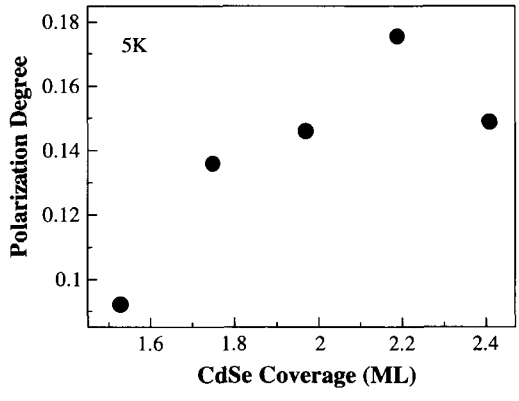


Fig. 6 CdSe coverage dependence of PL polarization degree at 5 K. (From Ref. [28].)

OPTICAL PROPERTIES OF CdSe QUANTUM DOTS

Microphotoluminescence

The FWHM of CdSe QDs is about several tens of microelectron volts, which is much broader than that of ZnCdSe quantum well.[29] The broad spectral width is considered to be a result of the nonuniformity of CdSe QD structures. The fluctuation in dot size causes the fluctuation in strain, resulting in more nonuniform dot structures and the broad spectral width. So it is impossible to clarify the optical properties of the single CdSe QD by macro-PL. To reduce the number of observed QDs, we

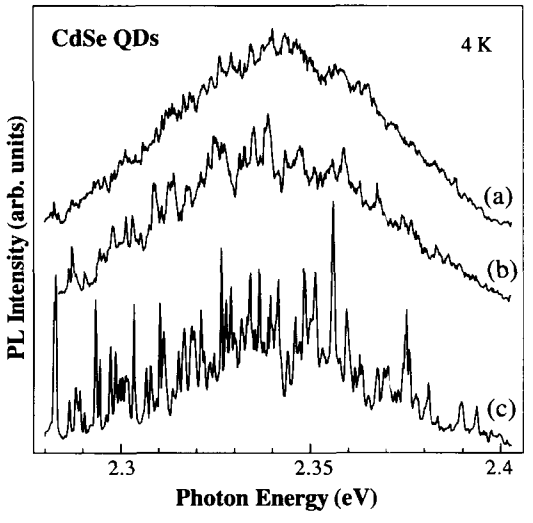


Fig. 7 Microphotoluminescence spectra of CdSe QDs at 4 K: (a) without Al mask, (b) 1 μm, and (c) 0.6-μm size aperture. (From Ref. [24].)

put Al masks with submicron size apertures on the sample surfaces.[30] The 100-nm-thick Al masks were made by electron-beam lithography and liftoff techniques. The diameters of the apertures were varied from 0.6 to 1 μm. Microphotoluminescence measurements were performed as follows. The sample fixed on a cold finger in a cryostat was cooled down using helium gas. The 458-nm line of Ar-ion laser was used to excite the QDs. The excitation laser was focused to about 0.8-μm spot size through a microscope objective. The CCD camera cooled by liquid nitrogen detected the luminescence through a monochromator. The spectral resolution is ~ 500 μeV. The micro-PL experiments were performed in the temperature range of 4 to 60 K.

Fig. 7 shows micro-PL spectra of CdSe QDs, which are measured through several different size apertures at 4 K. The excitation power of each spectrum was 100 μW. Fig. 7a shows a spectrum of micro-PL with focusing the excitation laser on the sample surface without Al masks. The wide FWHM of the broad peak is a result of the size distribution of the QDs. With decreasing the size of the aperture from 1 to 0.6 μm, as shown in Fig. 7b and c, respectively, the broad peaks observed through the apertures split into a number of anomalously very sharp lines having linewidths of ~ 500 μeV. These sharp lines in Fig. 7c indicate zero-dimensional density of states of the QDs. The characteristic features of the spectrum in Fig. 7c are as follows. In lower-energy region ranged from 2.28 to 2.34 eV, the sharp lines are clustered close together in comparison with those in the higher-energy region ranged from 2.34 to 2.40 eV. The integrated intensities of the lower-energy lines are larger than those of the higher-energy lines. The linewidths of the lower-energy lines are narrower than those of the higher-energy lines. These phenomena are similar to the result in the report of CdSe QDs by Kim et al.[31] They also found that the micro-PL spectra resulted from two different kinds of states. The micro-PL spectrum in Fig. 7c seems to have the broad band under the sharp lines. However, this is not clear at the present stage.

Fig. 8 shows the temperature dependence of micro-PL spectra measured through the 0.6-μm aperture. The temperature was ranged from 4 to 60 K. The excitation power is 100 μW. Fig. 8a and b is obtained from lower- and higher-energy regions, respectively. These spectra reveal that the peaks have Lorenzian line shapes. The linewidths shown in Fig. 8a, which come from larger QDs, are almost constant until 20 K and increase rapidly above 30 K. At 60 K, they are ~ 3.5 meV, which are about six times larger than that at 4 K. The behavior shown in Fig. 8a is similar to the result in the report of InAs QDs.[32] With increasing the temperature, excitons in the ground state are gradually scattered into higher states by the absorption of acoustic phonons. The increase

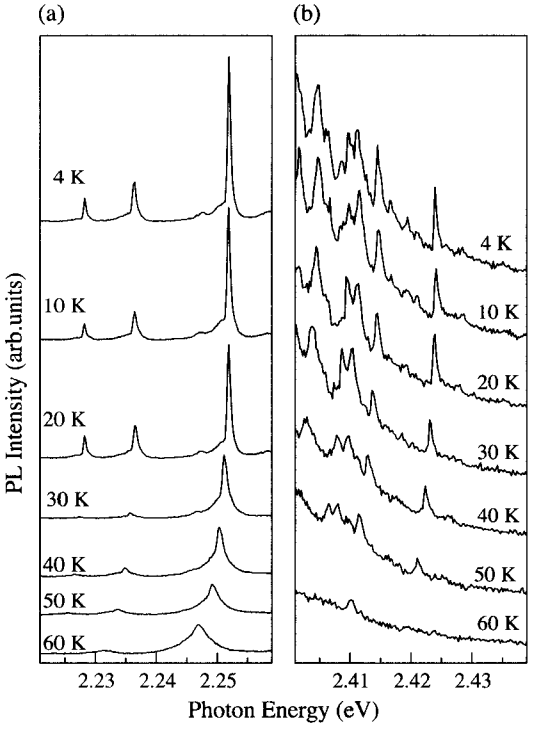


Fig. 8 Temperature dependence of micro-PL spectra of CdSe QDs taken through the 0.6-μm aperture. (From Ref. [24].)

in the FWHMs with temperature shown in Fig. 8a is a result of the increase of the scattering rate. On the other hand, the linewidths shown in Fig. 8b, which come from smaller QDs, hardly increase with the temperature. These results reveal that the degree of the linewidth broadening depends on the peak energy position, corresponding to the QD size, and that the linewidths of the smaller QDs are less temperature-dependent than those of the larger QDs. The difference shown in Fig. 8a and b might be able to be explained with the use of the energy level spacing of CdSe QDs. The smaller QDs have larger energy level spacing than the larger QDs because the smaller QDs have larger confinement energy, resulting in the reduction in the scattering rate of excitons from the ground state to higher states. As a result, the linewidths of smaller QDs are considered to be less temperature-dependent, as shown in Fig. 8b.

Selectively Excited Photoluminescence

A 500-W Xe lamplight dispersed with a 0.3-m monochromator served as a tunable light source for selectively excited PL measurements. The light was focused on the sample with rectangle shape using a microscope objective. The spectral width of the dispersed light was about 0.3 nm at 500 nm. The sample was mounted on a cold

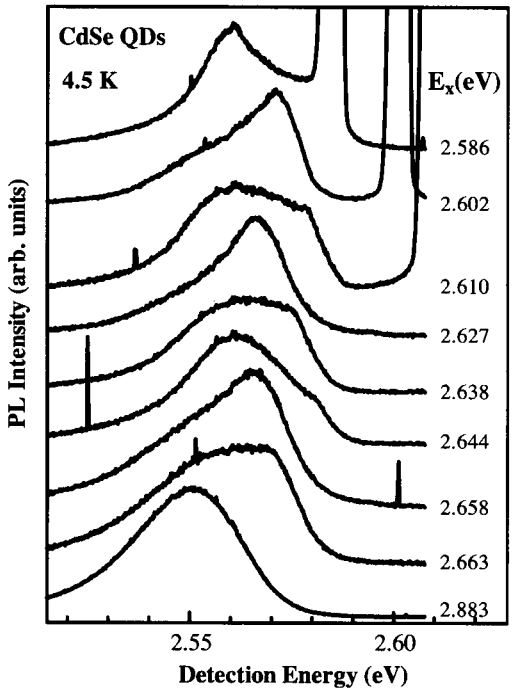


Fig. 9 Photoluminescence spectra of 1.6 ML CdSe coverage sample at different excitation energies at 4.5 K.

finger in vacuum in a small, continuous gas flow cryostat. The luminescence was collected using optical glass fibers, dispersed with a 0.5-m single-grating monochromator, and detected with a CCD camera cooled by liquid nitrogen. Fig. 9 shows the PL spectra of CdSe QDs at different excitation energy at 4.5 K. The spectrum at the bottom of Fig. 9 is measured for excitation above the band-gap energy of the ZnSe barrier layers. In this spectrum, there is a peak at 2.551 eV with a FWHM of 30 meV, which originates from CdSe QDs.[27] The shape of this spectrum is almost the same as that excited by a He–Cd laser. The broad spectral width is a result of the inhomogeneous size of CdSe QDs.[24,28] When the excitation energy decreases, the PL intensities are dramatically reduced. In addition, the spectra have a number of shape features at higher detection energies, as shown in Fig. 9. These peak positions depend upon the excitation energy. However, the features of the PL spectra in Fig. 9 do not seem to change regularly with decreasing excitation energy. These results are considered to be due to the selective excitation of QDs with different sizes in their ground states or in their excited states.

To clarify the origin of the spectra in Fig. 9, the PL spectra of CdSe QDs are measured successively in the excitation energy range from 2.82 to 2.55 eV at 4.5 K. Fig. 10 shows a contour map of the PL intensity of CdSe QDs as a function of the detection energy and the excitation energy. Energy scale of the abscissa is expanded as compared with that of the ordinate. Horizontal sections at several excitation energies correspond to the PL spectra in Fig. 9. The abrupt and quite strong absorption edge is clearly observed at 2.81-eV excitation energy for whole detection energy, which corresponds to the band-gap energy of the ZnSe barrier layer. This reveals that carriers generated in the ZnSe barrier layers diffuse in the sample efficiently into CdSe QDs and then emit luminescence there. This result indicates that CdSe QDs with high quality are formed on ZnSe surfaces.

At lower PL detection energy below approximately 2.54 eV in Fig. 10, a quite weak absorption band is observed at about 2.60-eV excitation energy, which is independent of detection energy, and a step-like broad absorption band above 2.675-eV excitation energy is also found, which continues up to ZnSe barrier absorption edge at 2.81-eV excitation energy. Moreover, several sharp absorption structures appear at higher PL detection energy, as shown in Fig. 10. The energy differences of these peaks are found to be roughly independent of the detection energy. These energy separations are estimated to be about 30 meV, which corresponds to the longitudinal-optical (LO) phonon energy in CdSe/ZnSe systems. Particularly, multiple LO-phonon structures up to the 6th order are clearly observed at around 2.58-eV detection energy. These phonon structures are terminated at the band-gap energy of ZnSe barrier layer. As compared with the PL excitation spectra of ZnCdSe reference quantum wells, absorption peaks from heavy- and light-hole excitons are not observed in that of CdSe QDs. These results indicate that the excitons in CdSe QDs are relaxed by multiple LO-phonon processes, resulting from strong confinement of excitons in zero-dimensional structures. Above the ZnSe band-gap energy, phonon structures in PL spectra are not observed, as shown in the PL spectrum at

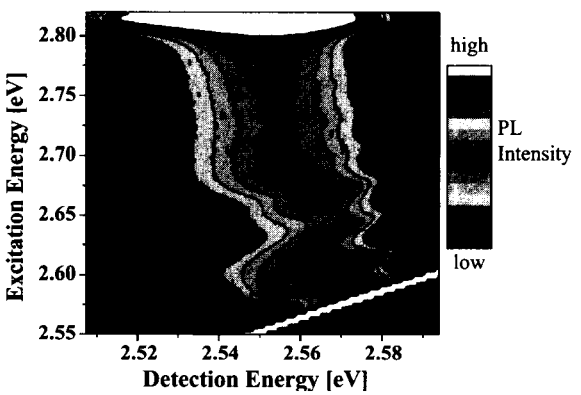


Fig. 10 Contour map of PL intensity of CdSe QDs at 4.5 K as a function of detection energy and excitation energy. (*View this art in color at www.dekker.com.*)

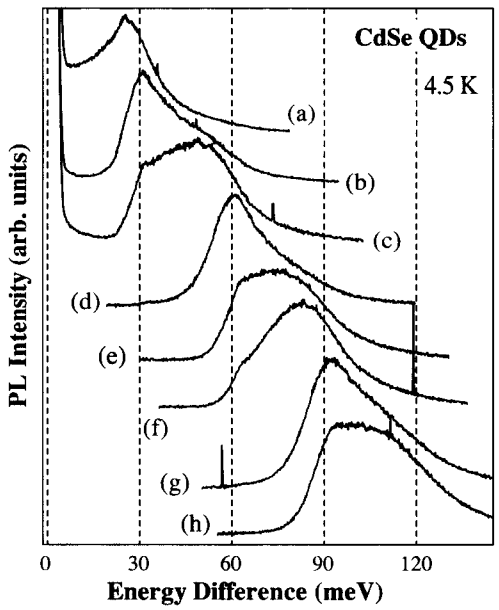


Fig. 11 Photoluminescence spectra of CdSe QDs at 4.5 K as a function of the energy difference between the excitation and detection energies. These PL spectra are the same as those in Fig. 9. Photoluminescence spectrum at the bottom in Fig. 9 is not used in this figure.

the bottom of Fig. 9, which is a result of the continuum states in the ZnSe layers.

To investigate the details of the relaxation process, we rearrange the PL spectra of CdSe QDs at 4.5 K as a function of the energy differences between the excitation and detection energies, as shown in Fig. 11. These spectra are obtained by horizontally shifting the PL spectra in Fig. 9 by the amount corresponding to the excitation energies. Although, as mentioned above, the spectra have the peaks at around 30, 60, and 90 meV, the peak positions are slightly dependent on excitation energies, as shown in Fig. 11. A peak is found at 25 meV in spectrum (a) in Fig. 11. This energy agrees with the 1-LO-phonon energy of CdSe bulk. This indicates that pure CdSe region exists in QDs, which seems to be inconsistent with the results of TEM measurements.[33–35] The spectra (b), (c), and (d) in Fig. 11 have peaks at 31 meV, and the spectra (e), (f), and (g) in Fig. 11 have peaks at 62 meV, corresponding to the 1-LO- and 2-LO-phonon energies of ZnSe bulk, respectively.

The relationship between phonon energy and detection energy of phonon structures can be obtained from Figs. 9 and 11. Fig. 12 shows the 1-LO- and 2-LO-phonon energies as functions of the detection energy of the phonon structures in PL spectra, which are obtained from Fig. 10. For detection at lower-energy side, the observed 1-LO- and 2-LO-phonon energies are estimated to be 25 and 50 meV, respectively, which agree with the LO-phonon energies of CdSe bulk. For higher detection energy side, phonon replicas can be resolved with energies of 31 and 62 meV, which correspond to the LO-phonon energies of ZnSe bulk. The size of CdSe QDs at higher detection energy is smaller than that at lower detection energy. This means that the wave functions of excitons at lower detection energy are confined better inside CdSe QDs than those at higher detection energy. Therefore for detection at lower-energy side (larger QDs), excitons are more interacted with the LO phonons of CdSe QDs, resulting in relaxation by using CdSe LO phonons, as shown in Fig. 12. On the other hand, the excitons at higher detection energies (smaller QDs) are weakly localized and their wave functions extend to the surrounding ZnSe barrier layers. Therefore excitons are more interacted with LO phonons of ZnSe barrier layer. As a result, excitons are relaxed by using ZnSe LO phonons, as shown in Fig. 12. At middle detection energy region (from 2.56 to 2.67 eV), phonon energies gradually change from 25 to 31 meV. These phonon energies might be a result of the contributions of both ZnSe and CdSe phonons or of ZnCdSe mixed crystals. The investigations of the Raman spectra of ZnCdSe bulk mixed crystal have shown that the LO-phonon spectra vs. crystal composition are the one-mode type,[36] indicating that the LO-phonon energy has changed monotonically with composition. It is difficult to clarify that excitons are relaxed by using both CdSe and ZnSe LO phonons or ZnCdSe LO phonons at middle detection energy region.

Next, the effect of interactions between CdSe QDs will be discussed. Fig. 13a shows a PL spectrum excited at

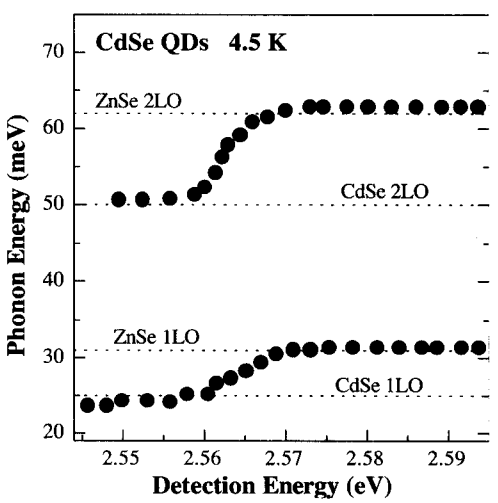


Fig. 12 1-LO- and 2-LO-phonon energies as functions of the detection energy of the phonon structures in PL spectra, which are based on the results shown in Fig. 10.

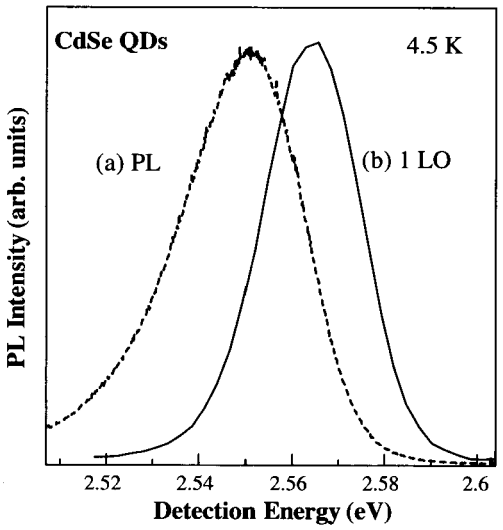


Fig. 13 (a) Photoluminescence spectrum excited at 2.833 eV above the ZnSe band-gap energy at 4.5 K and (b) a distribution curve of PL intensity related with the 1-LO-phonon structure obtained from Fig. 10.

2.833 eV for above ZnSe band-gap energy at 4.5 K, which has a peak at 2.551 eV. A distribution curve of PL intensity related with the 1-LO-phonon structure is also shown in Fig. 13b, which is obtained from Fig. 10. This curve has a peak at 2.565 eV. The distribution curve (b) in Fig. 13 shifts to higher detection energy by about 14 meV than the PL spectrum (a) in Fig. 13. In the case of InAs QDs on GaAs substrates, the intensity curves of LO phonons were identical to PL spectra for the above GaAs band-gap excitation.[30] This difference in the peak positions in Fig. 13a and b is considered to be a result of the interaction between CdSe QDs.

Because wave functions at higher excited states of CdSe QDs are more extended than those at the ground states, the interactions between the wave function of QDs become stronger. Then, when CdSe QDs are excited at higher excitation energy, carriers can easily move from smaller QDs to larger QDs. As a result, lower detection energy side in PL intensity is enhanced, as shown in the spectrum in Fig. 13a. The broad absorption band at 2.675-eV excitation shown in Fig. 10 is considered to result from the strong interactions between the wave functions of QDs. This broad absorption band is continuous with the ZnSe barrier absorption edge.

On the other hand, because wave functions are confined better at the ground states in CdSe QDs, the interactions between the wave function of QDs are quite weak. Then, when CdSe QDs are excited at lower excitation energy, PL is obtained from the same CdSe QDs. Consequently, the distribution curve in Fig. 13b is considered to represent real size distribution curve of CdSe QDs. However, the weak absorption, which is independent of detection energy, is obtained at 2.60-eV excitation energy in Fig. 10. This might result from the weak interaction between wave functions of QDs at the ground states.

Recently, carrier relaxation dynamics of CdSe QDs have been investigated by time-resolved PL measurements,[37] revealing that the decay time at higher-energy side is decreased with increasing detection energy. This result is considered to be due to the interactions between CdSe QDs, which is consistent with that of the selectively excited PL measurements.

CONCLUSION

We have investigated the formation and optical properties of self-organized CdSe QDs on ZnSe (001) surfaces with the use of TEM and PL measurements. CdSe QDs with high quality are naturally formed on ZnSe surfaces by MBE when the thickness of CdSe layers is less than about 2 ML. CdSe QDs show a strong green emission even at room temperature. The polarization dependences of PL from CdSe QDs exhibit that the PL spectra are strongly polarized parallel to the $[1\bar{1}0]$ direction. The CdSe coverage dependence of polarization PL spectra indicates that PL polarization degrees increase with QD formation.

Using micro-PL, single CdSe QD has high-resolved sharp lines, indicating zero-dimensional density of states of the QDs. It is found that the linewidths of the larger-size QDs are more temperature-dependent than those of smaller ones, suggesting to be a result of the difference in the energy level spacing of CdSe QDs.

The carrier relaxation mechanisms in self-organized CdSe QDs have been also investigated by selectively excited PL measurements. The contour map of PL intensity of CdSe QDs reveals that the strong and abrupt absorption edge of the ZnSe barrier layer is clearly observed for whole detection energy. In addition, photoexcited carriers that form excitons are relaxed by emitting multiple LO phonons in CdSe QDs, resulting from the strong confinement of excitons in zero-dimensional structures. At low and high detection energy, LO phonons of CdSe and ZnSe are dominated for relaxation process, respectively. The weak absorption and broad absorption bands observed at lower PL detection energy in the contour map can be interpreted by the interactions between wave functions of CdSe QDs. Stronger interactions between wave functions of CdSe QDs at higher excitation energy induce the difference

in the peak positions between PL spectrum in Fig. 13a and the distribution curve in Fig. 13b. Therefore the distribution curve in Fig. 13b is considered to represent the real size distribution curve of CdSe QDs.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Koichi Inoue, Dr. Takeshi Ota, Dr. Yasuhiro Murase, and Mr. Nobuhiro Yasui of Osaka University for experimental supports and useful discussions. The authors are also grateful to Prof. Yoshihiko Hirotsu, Dr. Takeshi Hanada, and Mr. Takeshi Ishibashi of Osaka University for TEM operations and Prof. Kazuo Murase and Dr. Kenichi Oto of Osaka University for electron-beam lithography and liftoff techniques.

REFERENCES

1. Freeman, M.R.; Awschalom, D.D. Femtosecond probes of magnetic phenomena: Spin-polarization spectroscopy in dilute-magnetic-semiconductor quantum wells. J. Appl. Phys. **1990**, *67*, 5102–5107.
2. Haase, M.; Qiu, J.; DePuydt, J.; Cheng, H. Blue-green laser diodes. Appl. Phys. Lett. **1991**, *59*, 1272–1274.
3. Ema, K.; Kuwata-Gonokami, M.; Shimizu, F. All-optical sub-Tbits/s serial-to-parallel conversion using excitonic giant nonlinearity. Appl. Phys. Lett. **1991**, *59*, 2799–2801.
4. Ding, J.; Jeon, H.; Ishihara, T.; Hagerott, M.; Nurmikko, A.V.; Luo, H.; Samarth, N.; Furdyna, J. Excitonic gain and laser emission in ZnSe-based quantum wells. Phys. Rev. Lett. **1992**, *69*, 1707–1710.
5. Bagnall, D.M.; Chen, Y.F.; Zhu, Z.; Yao, T.; Koyama, S.; Shen, M.Y.; Goto, T. Optically pumped lasing of ZnO at room temperature. Appl. Phys. Lett. **1997**, *70*, 2230–2232.
6. Esaki, L.; Tsu, R. Superlattice and negative differential conductivity in semiconductor. IBM J. Res. Develop. **1970**, *14*, 61–65.
7. Gershoni, D.; Temkin, H.; Dolan, G.J.; Dunsmuir, J.; Chu, S.N.G.; Panish, M.B. Effects of two-dimensional confinement on the optical properties of InGaAs/InP quantum wire structures. Appl. Phys. Lett. **1988**, *53*, 995–997.
8. Kapon, E.; Simhony, S.; Bhat, R.; Hwang, D.M. Single quantum wire semiconductor lasers. Appl. Phys. Lett. **1989**, *55*, 2715–2717.
9. Tsuchiya, M.; Gaines, J.M.; Yan, R.H.; Simes, R.J.; Holtz, P.O.; Coldren, L.A.; Petroff, P.M. Optical anisotropy in a quantum-well-wire array with two-dimensional quantum confinement. Phys. Rev. Lett. **1989**, *62*, 466–469.
10. Goni, A.R.; Pfeiffer, L.N.; West, K.W.; Pinczuk, A.; Baranger, H.U.; Stormer, H.L. Observation of quantum wire formation at intersecting quantum wells. Appl. Phys. Lett. **1992**, *61*, 1956–1958.
11. Nagamune, Y.; Tsukamoto, S.; Nishioka, M.; Arakawa, Y. Growth process and mechanism of nanometer-scale GaAs dot-structures using MOCVD selective growth. J. Cryst. Growth **1993**, *126*, 707–717.
12. Tsutui, K.; Hu, E.; Wilkinson, W. Reactive ion etched II–VI quantum dots: Dependence of etched profile on pattern geometry. Jpn. J. Appl. Phys. **1993**, *32*, 6233–6236.
13. Goldstein, L.; Glas, F.; Marzin, J.; Charasse, M.; LeRoux, G. Growth by molecular beam epitaxy and characterization of InAs/GaAs strained-layer superlattices. Appl. Phys. Lett. **1985**, *47*, 1099–1101.
14. Leonard, D.; Krishnamurthy, M.; Reaves, C.; DenBaars, S.; Petroff, P. Direct formation of quantum-sized dots from uniform coherent islands of InGaAs on GaAs surfaces. Appl. Phys. Lett. **1993**, *63*, 3203–3205.
15. Fafard, S.; Leonard, D.; Merz, J.L.; Petroff, P.M. Selective excitation of the photoluminescence and the energy levels of ultrasmall InGaAs/GaAs quantum dots. Appl. Phys. Lett. **1994**, *65*, 1388–1390.
16. Marzin, J.-Y.; Gerard, J.-M.; Izrael, A.; Barrier, D.; Bastard, G. Photoluminescence of single InAs quantum dots obtained by self-organized growth on GaAs. Phys. Rev. Lett. **1994**, *73*, 716–719.
17. Oshinowo, J.; Nishioka, M.; Ishida, S.; Arakawa, Y. Highly uniform InGaAs/GaAs quantum dots ($\sim$ 15 nm) by metalorganic chemical vapor deposition. Appl. Phys. Lett. **1994**, *65*, 1421–1423.
18. Grundmann, M.; Christen, J.; Ledentsov, N.; Böhrer, J.; Bimberg, D.; Ruvimov, S.; Werner, P.; Richter, U.; Gösele, U.; Heydenreich, J.; Usti-nov, V.; Egorov, A.; Zhukov, A.; Kop'ev, P.; Alferov, Zh. Ultranarrow luminescence lines from single quantum dots. Phys. Rev. Lett. **1995**, *74*, 4043–4046.
19. Steer, M.J.; Mowbray, D.J.; Tribe, W.R.; Skolnick, M.S.; Sturge, M.D.; Hopkinson, M.; Cullis, A.G.; Whitehouse, C.R. Electronic energy levels and energy relaxation mechanisms in self-organized InAs/GaAs quantum dots. Phys. Rev., B **1996**, *54*, 17738–17744.
20. Heitz, R.; Veit, M.; Ledentsov, N.N.; Hoffmann, A.;

Bimberg, D.; Ustinov, V.M.; Kop'ev, P.S.; Alferov, Zh.I. Energy relaxation by multiphonon processes in InAs/GaAs quantum dots. Phys. Rev., B **1997**, *56*, 10435–10445.

21. Maehashi, K.; Morota, N.; Murase, Y.; Nakashima, H. Molecular beam epitaxial growth of ZnSe films on vicinal GaAs(110) substrates. Jpn. J. Appl. Phys. **1999**, *38*, 1339–1342.
22. Maehashi, K.; Morota, N.; Murase, Y.; Yasui, N.; Shikimi, A.; Nakashima, H. Formation of the charge balanced ZnSe/GaAs(110) interfaces by molecular beam epitaxy. J. Cryst. Growth **1999**, *201/202*, 486–489.
23. Murase, Y.; Noma, T.; Maehashi, K.; Nakashima, H. Formation and photoluminescence of stacked CdSe quantum dots grown by molecular beam epitaxy. J. Vac. Sci. Technol., B **2001**, *19*, 1459–1462.
24. Maehashi, K.; Yasui, N.; Murase, Y.; Ota, T.; Noma, T.; Nakashima, H. Formation and characterization of self-organized CdSe quantum dots. J. Electron. Mater. **2000**, *29*, 542–548.
25. Kirmse, H.; Schneider, R.; Rabe, M.; Neumann, W.; Henneberger, F. Transmission electron microscopy investigation of structural properties of self-assembled CdSe/ZnSe quantum dots. Appl. Phys. Lett. **1998**, *72*, 1329–1331.
26. Xin, S.; Wang, P.; Yin, A.; Yin, C.; Dobrowolska, M.; Merz, J.; Furdya, J. Formation of self-assembling CdSe quantum dots on ZnSe by molecular beam epitaxy. Appl. Phys. Lett. **1996**, *69*, 3884–3886.
27. Maehashi, K.; Yasui, N.; Murase, Y.; Shikimi, A.; Nakashima, H. Formation of self-organized CdSe quantum dots on ZnSe(100) surfaces by molecular beam epitaxy. Appl. Surf. Sci. **2000**, *166*, 322–325.
28. Maehashi, K.; Yasui, N.; Ota, T.; Noma, T.; Murase, Y.; Nakashima, H. Structural and optical properties of CdSe/ZnSe self-organized quantum dots. J. Cryst. Growth **2001**, *227–228*, 1116–1120.
29. Murase, Y.; Ota, T.; Yasui, N.; Shikimi, A.; Noma, T.; Maehashi, K.; Nakashima, H. Temperature dependence of photoluminescence spectra of self-organized CdSe quantum dots. J. Cryst. Growth **2000**, *214/215*, 770–773.
30. Ota, T.; Murase, Y.; Noma, T.; Maehashi, K.; Nakashima, H.; Oto, K.; Murase. K. Micro-photoluminescence from CdSe quantum dots. J. Cryst. Growth **2000**, *214/215*, 778–781.
31. Kim, J.C.; Rho, H.; Smith, L.; Jackson, H.E.; Lee, S.; Dobrowolska, M.; Furdyna, J.K. Temperature-dependent micro-photoluminescence of individual CdSe self-assembled quantum dots. Appl. Phys. Lett. **1999**, *75*, 214–216.
32. Ota, K.; Usami, N.; Shiraki, Y. Temperature dependence of microscopic photoluminescence spectra of quantum dots and quantum wells. Physica, E **1998**, *2*, 573–577.
33. Strassburg, M.; Kutzer, V.; Pohl, U.W.; Hoffmann, A.; Broser, I.; Ledentsov, N.N.; Bimberg, D.; Rosenauer, A.; Fischer, U.; Gerthsen, D.; Krestnikov, I.L.; Maximov, M.V.; Kop'ev, P.S.; Alferov, Zh.I. Gain studies of (Cd, Zn)Se quantum islands in a ZnSe matrix. Appl. Phys. Lett. **1998**, *73*, 942–944.
34. Kümmell, T.; Weigand, R.; Bacher, G.; Forchel, A.; Leonardi, K.; Hommel, D.; Selke, H. Single zero-dimensional excitons in CdSe/ZnSe nanostructures. Appl. Phys. Lett. **1998**, *73*, 3105–3107.
35. Litvinov, D.; Rosenauer, A.; Gerthsen, D.; Ledentsov, N.N. Character of the Cd distribution in ultrathin CdSe layers in a ZnSe matrix. Phys. Rev., B **2000**, *61*, 16819–16826.
36. Lozykowski, H.J.; Shastri, V.K. Excitonic and Raman properties of ZnSe/$Zn_{1-x}Cd_xSe$ strained-layer quantum wells. J. Appl. Phys. **1991**, *69*, 3235–3242.
37. Ota, T.; Murase, Y.; Maehashi, K.; Nakashima, H.; Watatani, C.; Edamatsu, K.; Itoh, T.; Oto, K.; Murase, K. Investigation of carrier relaxation dynamics in single CdSe/ZnSe self-organized quantum dot by time-resolved micro-photoluminescence. J. Electron. Mater. **2001**, *30*, 448–452.

Quantum Dots Made of Metals: Preparation and Characterization

Q

J. P. Wilcoxon
Sandia National Laboratories, Albuquerque, New Mexico, U.S.A.

INTRODUCTION

Nanosize metal clusters occupy a position between the molecular and solid state and, because of the dominant role of their abundant surface area, provide a unique way to learn how metal–metal bonding, cluster shape, and packing are affected by ligands bound to the cluster surface. Such studies may give insights into complex issues in catalysis, such as selectivity of binding of substrates to vertex, edge, or face sites on a metal cluster, and how such binding affects the intermetal bond distances, essentially causing a surface reconstruction, mass redistribution, or shape change. Studies of the 3-D interface structure of nanosize metal clusters should yield information quite different from the extensive literature describing ligand interactions with extended, 2-D metal surfaces. Because clusters more closely resemble practical heterogeneous catalysts (e.g., hydrodesulfurization process, hydrogenation process), important new scientific and technical insights may be gained by their study.

The formation of metal colloids or clusters by the controlled reduction, nucleation, and growth from metal salts in aqueous solution has been investigated for over a century. In the case of gold colloids, the earliest scientific investigations were undertaken by Michael Faraday[1] and spanned over a decade. Human interest in the colors of such colloidal systems dates back to the Roman times, as pointed out by Kreibig and Vollmer[2] in their excellent review of the optical properties of metal clusters. An example cited therein is a Roman goblet from the fourth century in the British Museum whose fame is attributed to the shining colors generated by a composition of Ag and Au clusters. The purple colors of colloidal Au and Cu dispersions are thought to be the origin of the association of purple with royal colors.

OVERVIEW

The earliest methods of colloidal synthesis (e.g., Faraday method) rely on the chemical reduction of metal salts in aqueous solution. The remarkable stability of these ancient wine-red sols was a result of charge stabilization via adsorbed citrate ions. In sealed vials, samples prepared by Faraday in the mid-1850s can be still be viewed in the Cavendish museum in Cambridge—a true testament to solution stability.

Only very recently have scientists developed nonaqueous methods of colloidal synthesis requiring steric, not charge, stabilization. An early example of this approach, which ultimately led to the development of magnetic particles for recording purposes, was that of Hoon et al.[3] They produced Co colloids (now called nanoclusters, nanocrystals, or nanoparticles) in the 1–100 nm range by thermolysis of $Co_2(CO)_8$ in the presence of dispersant polymers. This seminal work provided the basis of nearly all subsequent methods for production of colloids from metallorganic precursors using high T decomposition. Other reports of base metal colloidal synthesis via thermal decomposition in the presence of surfactants or polymers soon followed.[4,5] With the advent of exotic organometallic precursors, even binary semiconductors could be produced by this approach.[6]

Although organometallic precursors are the basis of many current approaches to cluster synthesis in nonaqueous, low-dielectric constant solvents, there are several drawbacks to this approach. The toxicity, air sensitivity, cost, and unwanted by-products (e.g., metallic films) prevented the widespread adoption of such methods outside the chemical community.

Methods of systematic control of cluster size in organometallic decomposition routes are still lacking. Nevertheless, in certain cases, remarkable monodispersity can be achieved. One of the earliest and best-known examples of monodisperse metal cluster growth from atomic precursors in solution was the synthesis by Schmid[7] of Au(N=55) clusters by diborane reduction of $Au(PPh_3)Cl$ in benzene or methylene chloride. Fifty-five atoms of a metal corresponds to the closing of the second atomic shell of a Au(N=13) cluster core with 42 Au atoms with a dozen PPh_3 molecules weakly ligated to the faces of the inorganic core. Consequently, the role of the ligand used to sterically stabilize inorganic nanoclusters in a solvent was discovered to be critical in determining the eventual structure and stable size of a given nanocluster. By a continuation of this growth process, Schmid has extended the range of sizes available to many thousands of atoms, utilizing a synthesis

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120009132

reminiscent of polymeric growth—one generation providing the "seeds" for the next.[8] His approach has also been extended to other metals such as Pd and Pt. However, the yields from each growth generation are low so, as in organic synthesis, the overall yield can be quite small for larger clusters. Also, the weak binding of the PPh3 group means that the long-term stability (e.g., longer than a few days) of such clusters is severely compromised.

It would be very nice to have a general nanocluster chemical synthesis, which combines the advantages of low toxicity/cost precursors, and high yield of traditional aqueous-based colloidal chemistry with the size dispersion control and chemical versatility of organometallic methods. Possibly with this goal in mind, Boutonnet et al.[9] first used aqueous pools of water found in oil—continuous microemulsions to solubilize simple, ionic, metal salts of Au, Pd, Pt, and Rh, followed by chemical reduction using hydrazine or hydrogen gas to produce metal clusters dispersed in oils. Such new nanomaterials were later shown to have good catalytic activity for hydrogenation.[10]

A serious limitation of using water containing microemulsions as microscopic reactors for metal colloid formation was the limited types of reducing agents permitted (basically hydrazine or aqueous $NaBH_4$), as well as the inherent polydispersity of microemulsions, which produced a fairly broad size dispersion in the final product. This limitation was overcome by the surprising discovery that ionic metal salts could be directly solubilized in a variety of newly discovered inverse micelle systems.[11–13] The most useful of these systems were reported in a series of papers and a patent, and form the basis of the most versatile method for the formation of nanoclusters, both semiconductor and metallic, in the 1–10 nm regime.[14–16] The complete absence of water and air in these systems allowed the use of very strong, even pyrophoric, reductants such as $LiAlH_4$ [in tetrahydrofuran (THF)] for the first time and permitted even exotic nanocrystalline species such as Si and Ge to be made by low-T chemical synthesis for the first time.[17–20] In addition, the well-defined interface and monodispersity inherent to inverse micelle systems resulted in narrow cluster size distributions.

Meanwhile, physicists were taking a different approach to cluster synthesis, based on their expertise in high-vacuum, molecular beam techniques coupled with the use of mass spectrometers for mass selection and analysis. An excellent example of the results that emerged from such endeavors is that of de Heer and coworkers, who utilized a supersonic expansion of an atomic cluster source into a vacuum in the presence of an inert gas to create cluster beams.[21] Using a mass spectrometer they, and others,[22] discovered that certain masses of clusters were produced in relatively large abundance. These "magic" sizes corresponded to the closing of atomic shells, analogous to stable nuclear shells from nuclear chemistry,[23] or the electronic shells forming the basis of chemical bonding and the periodic table. Alkali metal clusters, in particular, were discovered to be well described by simple "jellium" models and their optical properties were investigated. Other interesting phenomena emerged from these studies, including the discovery of enhanced magnetic moments in small clusters of Co, Fe, and Ni compared to their bulk counterparts.[24]

The ability to compare the size-dependent physical properties (e.g., absorbance, PL, magnetic behavior, melting point, chemical reactivity) of clusters prepared by beam methods and those prepared by solution techniques is complicated by the very important role of the interface, also known as the embedding media effect.[2] "Naked" clusters in a vacuum have very different optoelectronic properties compared to those coordinated to a solvent or a ligand. Importantly, clusters in solution, in order to be perfectly dispersed, require tightly binding ligands to prevent association or "clumping" in solution, so one can never consider very small inorganic nanoclusters (1–3 nm) without also understanding their organic interface or "shell." Even in a vacuum, a naked cluster has unsatisfied bonding at its surface, which probably requires some sort of structural "reconstruction" to minimize its energy, and this has not been taken into account in even the most refined theories of cluster structure. Thus, there is considerable controversy among theorists regarding the actual equilibrium shape of a cluster for a given number of atoms. In short, we have a long way to go before understanding the size- and interface-dependent properties of nanoclusters. Scientists may expect to be fruitfully employed in this enterprise for many years to come.

In this review, we cannot possibly do justice to the enormous literature in the area of metallic nanoclusters or quantum dots. Accordingly, we primarily focus on metallic Q-dots of Au and Ag, which historically provided the earliest fascination with dispersed colloidal materials. In fact, the beautiful colors of colloidal sols, which differed so significantly from their bulk counterparts and organic dyes, influenced their medieval uses in stained glass, goblets, etc. Their synthetic preparation was a well-guarded alchemic secret. It is probably not surprising that the majority of scientific investigations of the optical properties of metal clusters have focused on the few metals with visible absorbance. This restricts studies to the alkalis and noble metals (Au, Ag, Cu). This is because strong collective oscillations (plasmon resonances) require quasi-free electrons, as well as weak damping of these oscillations in the visible regime [i.e., the imaginary part of the dielectric constant $\varepsilon(\omega)$ must be small]. A further preference for studies of Au and Ag colloids occurs because of their remarkable long-term stabilities in aqueous solutions under ambient conditions and their

interesting size- and shape-dependent optical properties, which can often be described by the Mie theory[25] of electromagnetic scattering. In this review, we emphasize the important role of synthesis methods, advanced characterization tools for both feedback and understanding of size/optical properties, and formation of arrays of these dots. We discuss both alloys and core–shell nanoparticles of these materials. Only briefly discussed will be the magnetic and catalytic properties of nanoclusters.

EXPERIMENTAL

Nanocluster Synthetic Methods

Inverse micelle method

The inverse micelle synthesis has been extensively described in a series of papers[11,13–20,26–28] and patent.[12] A key difference of this method from either liquid or gas atomic aggregation processes (e.g., high-T thermal decomposition of organometallics) is that the cluster–cluster aggregation process, whose kinetically determined structures inevitably result in power-law or log normal cluster size distributions, is fundamentally altered by the microheterogeneous environment of the droplet-like inverse micelles. Because the ionic metal salt precursor is only soluble inside the micelle at typical precursor metal salt concentration of 0.01–0.1 M and surfactant concentrations of ~5–10 wt.% (~0.2 M), there are only about 1–4 precursor ions/micelle. Thus, growth to the final observed sizes of N=10–10000 atoms must occur via micellar diffusion, micelle collision, temporary interface fusion, and atomic interchange. The micelle interior volume roughly determines the maximum size of the small atomic clusters, which may interchange during the growth or aggregation step of the synthesis. This cluster growth rate is set by the diffusion rate of the micelles, which is ~2 orders of magnitude slower than would occur in a continuous liquid phase (i.e., Schmid approach, gas-phase aggregation, etc.). So the opportunity for structural adjustments to atomic positions during atomic exchange and growth is enhanced, possibly favoring thermodynamic stabilities over metastable (e.g., disordered) cluster structures. For this reason, we have never observed disordered, and rarely[29] nonequilibrium nanoclusters as the final product regardless of the melting point of the corresponding bulk material.

Our approach has changed only in minor ways since our first description of metal nanocluster synthesis in 1989.[11] The most significant change is the use of strongly binding surfactants, thiols, to passivate the surface during or after the chemical reduction, which allowed us to employ liquid chromatographic analysis of the nanocluster size and size dispersion.[30] When thiols are added prereduction, they can significantly alter the final nanocluster size, because they inhibit the cluster growth more strongly than the nonionic and cationic surfactants we typically use to solubilize ionic gold, silver, or platinum salts in inverse micelles. However, the use of alkyl thiols as passivating agents presents other issues, which we have recently discussed—certain alkyl thiols act as etchants, reducing the size of the as-synthesized clusters.[31] They also fail to bind as effectively to larger (e.g., D_c>4 nm) (D_c=inorganic core diameter) Au clusters compared to smaller ones (e.g., D_c<2–3 nm). The thiols, although present in much lesser amounts than the surfactants used to form the inverse micelles, compete very effectively for binding sites on the growing nanocluster surface. Because of their strong binding properties, thiol passivating agents also permit purification and removal of ionic byproducts and most of the surfactant micelles used to solubilize the metal salt precursors. The stabilization of the nanocluster surface by a strongly binding ligand is critical to the variety of purification approaches we have described.[30]

Numerous other papers describing Au nanocluster synthesis using so-called phase transfer catalysts (i.e., cationic surfactants)[32,33] have been published. There are some misconceptions about the role of the cationic surfactants used in this synthetic approach. The first is that the typical recipe requires water to first dissolve the salt (typically $HAuCl_4$ or $NaAuCl_4$) and then requires this aqueous salt solution to be brought into contact with an immiscible toluene solution containing a cationic surfactant [usually tetraoctylammonium bromide (TOAB)]. One needs to ask why only certain cationic surfactants will actually solubilize the gold salt into the organic phase? The reason is that very hydrophobic, long-alkyl chain surfactants are required to ensure both the formation of spherical micelles in the organic (toluene) phase and the total exclusion of water from that phase. Because such cationic surfactants spontaneously form inverse micelles in toluene, as was discovered several years ago by neutron scattering, they are capable of directly solubilizing a wide variety of metal salts without the use of water. We have described such an inverse micelle synthesis in Ref. [13] and in our original patent,[12] using didodecyldimethyl ammonium bromide (DDAB) in toluene. One can also use linear alkane solvents with these surfactants, but geometric considerations in the tail group packing often require a cosurfactant such as hexanol be added to "fill in the gaps" in the micellar surfactant droplet interface with linear alkanes.

With this general understanding of the role of the cationic surfactant as a micelle forming agent, which can directly solubilize a variety of metal salts, water need not

be introduced into the system, thus allowing the use of strong reducing agents such as $LiBH_4$ in tetrahydrofuran (THF), or $LiAlH_4$ in THF or toluene. These reducing agents are generally superior to $NaBH_4$ in water, which must be used to reduce $HAuCl_4$ in two-phase systems. Even if one chooses to use $NaBH_4$ as a reductant, because $NaBH_4$ in neutral aqueous solution is unstable with respect to hydrolysis, only highly alkaline solutions ($\sim$4 NaOH:$NaBH_4$) should be used to effect reductions. Such caustic reducing solutions are quite stable and are actually sold by Aldrich as stock $\sim$4.4 M $NaBH_4$ in $\sim$14 M NaOH solutions, and can be diluted into water, as desired.

In the case of gold or silver, the precursor metal salts are so easily reduced that the use of alkaline $NaBH_4$ in water as we described in our early work is quite acceptable. However, this reducing agent is somewhat ineffective for the reduction of Pt or Pd salts, and is completely ineffective for formation of metallic Fe, Ni, Co, Si, or Ge. Instead, in the case of Fe, Ni, or Co, the corresponding metal boride will form.

Metathesis method

In this cluster formation method, a chemical reduction of an organic soluble metal–organic precursor is used to produced metal clusters. This can be carried out in a coordinating or noncoordinating solvent, or some combination thereof. Alternatively, a surface-active (but not micelle forming) species such as dodecanthiol (C12SH) or tri-ocytylphosphine (TOP) may be added to prevent runaway growth during the chemical reduction. The best example of this approach is the one described by Schmid, and exemplified by the reaction that produces Au(N=55) clusters. We refer the reader to his excellent review on this subject for further details.[8]

Another approach to the production of metal clusters in solution relies on the high-T reduction of organometallic precursors by glycols in coordinating solvents such as diphenyl ether. This ''poly-ol'' method was recently refined to produce base metal clusters of a variety of sizes of Co, Fe, and Ni. Generally, empirically chosen surfactants are used to stabilize the growing nanocluster. A review of this approach appeared recently.[34] Briefly, the nanocluster size is increased by increasing the amount of precursor and/or decreasing the amount of surfactant. There is usually some unwanted metallic film formation and/or incomplete reduction so the yield is not 100%, and the method requires high boiling point solvents as the reduction typically takes place between 200 and 300°C. Some of these solvents may not allow solubilization of all the available components (metal–organic, poly-ol, stabilizer), so a fair amount of empirical variation is required before achieving an optimal synthesis.

Thermal decomposition method

In this method, which is the oldest approach for nonaqueous synthesis of metallic Q-dots, a thermally labile, oil soluble, metal–organic precursor is thermally decomposed in the presence of a surfactant-like stabilizer, typically a block copolymer.[3] Under the proper conditions, fairly monodisperse colloids can result. However, the method is typically only used for base-metal nanocluster synthesis (Co, Fe, Ni) because an available thermally unstable metallorganic is required [e.g., $Fe(CO)_5$, $Co_2(CO)_8$, $Ni(CO)_4$]. As these compounds are very air-sensitive and either mildly or severely toxic, all handling must be carried out in either a glove box, or by Schlenk-line methods. Additionally, because of the rapid growth rate compared to chemical reduction methods, the final clusters are highly defective, requiring significant annealing to produce high-quality samples. This means the final material properties (e.g., saturated magnetic response) are typically only a small fraction of that found in the corresponding bulk materials.[3–5]

Nanocluster Characterization

Size exclusion chromatography (SEC)

We use a commercial Water's corporation autosampler, solvent pump, degasser, and detectors together with commercially available polysytrene microgel columns (Polymer Laboratories) with controlled pore sizes to separate and analyze organically passivated metal nanoclusters. The alkyl organic shell on the metal surface prevents specific chemical interaction with the hydrophobic column material. A single-channel output of an in-line photodiode array (PDA) detector at a particular wavelength near the classical plasmon resonance, ($\sim$520 nm for Au, $\sim$400 nm for Ag) was used as described in previous papers[30,35] to obtain the peak apex elution time and peak width, which together yield the average cluster size and size dispersion. A conductivity detector and refractive index (RI) detector were also used to demonstrate that the eluting metal clusters had no charge and were separated from the nonabsorbing chemicals (surfactant micelles, salt byproducts, etc.) used in the synthesis. A chromatogram showing the separation D_c=2 nm Au nanoclusters from the solvents and surfactants used to produce them is illustrated in Fig. 1, taken from Ref. [30].

The use of SEC to obtain cluster sizes and size distributions was explained in detail in a recent paper.[30] In that work, we demonstrated that the elution time t_e for a properly passivated Au nanocluster sample with a hydrodynamic diameter D_h obeyed the relation, log $D_h \sim t_e$, allowing one to obtain a metal core diameter after subtraction of the thickness of the organic passivating

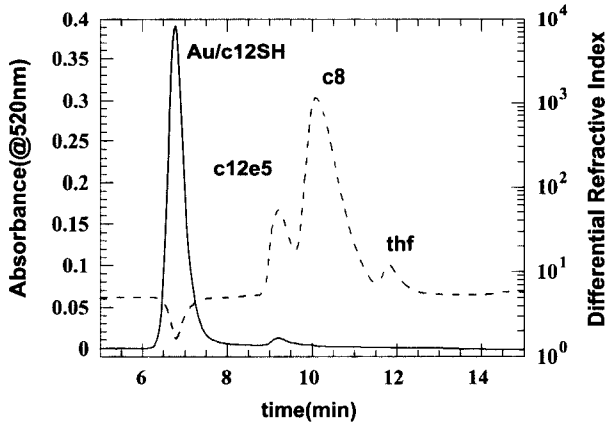


Fig. 1 Chromatographic separation of capped Au nanoclusters stabilized using c12SH and containing the nonionic surfactant c12e5. A PL500 column and a toluene mobile phase were used.

layer. This thickness was determined by taking a sample with a single core size as determined by transmission electron microscopy (TEM) and adding a series of alkyl thiols, designated herein as CkSH, with chain lengths k of $6<k<16$, to the toluene mobile phase, and obtaining D_h from t_e and subtracting the known TEM core size. Fig. 2 shows an example of this size separation capability for a single core size, $D_c=2.0$ nm and $k=6$, 10, and 14. It illustrates that clusters with different total hydrodynamic sizes due to organic shell differences may be separated by using SEC.

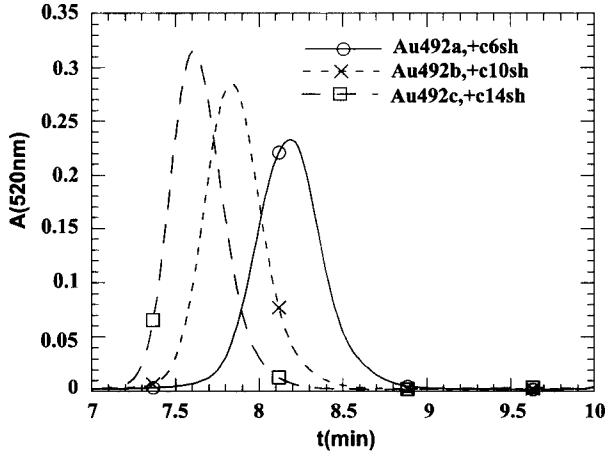


Fig. 2 Effect of capping agents on the hydrodynamic diameter of Au nanoclusters. The HPLC chromatogram of Au nanoclusters with three alkyl thiols, CkSH, $k=6$, 10, 14, added after reduction. The absorbance at 520 nm [A(520 nm)] from the PDA vs. elution time is shown. The SEC column used was a Polymer Labs PL1000 and the mobile phase was toluene at 1 mL/min flow rate. (From Ref. [30].)

The organic passivating shell thicknesses agree well with values calculated from known C—C bond lengths in linear alkanes and confirmed by SEC. A best fit to a series of linear alkanes and polysytrene polymer standards gave D_h (nm)=608 exp($-0.62525t_e$ (min)) for the PL1000 column used in that reference. This relation was valid for 1 nm $< D_h <$ 10 nm. Columns are available with larger pore sizes, which extend this range to larger values. For further information, the reader should consult the excellent book by Yau et al.[37]

In the case of other nanoclusters such as Si, Ge, Co, Ni, and Fe, or semiconductors such as CdS, CdSe, FeS_2, and MoS_2, tetrahydrofuran (THF) or acetonitrile (ACN) appears to be the best mobile phase choice. It is highly recommended to use inhibited THF to minimize oxidation as a result of peroxide formation. This is especially true if fraction collection of purified nanoclusters is desired. In our case, automatic vacuum removal of all gases occurs prior to the pump and injection system, so air-sensitive nanomaterials can be easily analyzed or fraction collected without degradation. In addition, an in-line filter and short "guard" column of the same column material is used to prevent retention time changes with age and/or injection of "bad" samples. With such an elaborate system (and many ruined columns during the learning process!), we can reproduce the retention time (i.e., D_h) of high-quality, stable nanocluster samples to within the instrument resolution of 0.01–0.02 min over periods exceeding 1 year. This corresponds to a size resolution (see below) of ~1 Å! Additionally, shape information is incorporated into the retention time, as we can achieve near baseline separation of C60 from C70 on a typical PL500 column, although these fullerenes differ in hydrodynamic size by only ~1 Å. Fig. 3 illustrates this point.

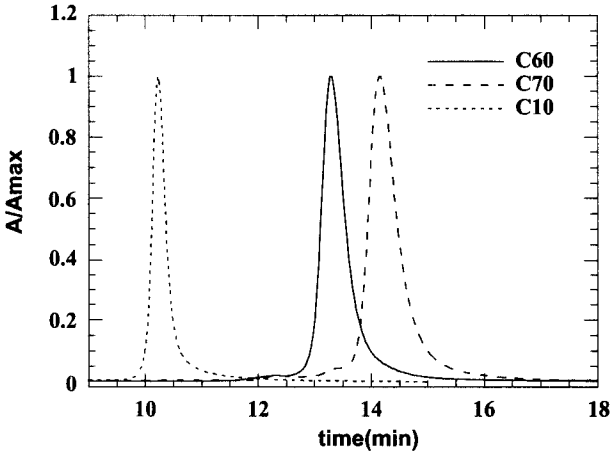


Fig. 3 Chromatogram instrumental linewidth for samples of decane (C10), C60, and C70 run on a PL500 column in toluene. (From Ref. [30].)

As we have discussed,[30,36] increasing the total pore volume for a pure SEC elution mechanism increases the resolving power of the separation. Preparative columns with diameters of 25.4mm are available (i.e., $\sim 10\times$ the pore volume) that can separate clusters differing in size by only $\sim 2\%$. With such columns, large amounts (10–100 mg) of nanoclusters can be injected and fraction collected with great size selectivity ($\sim$1–2 Å separation). We have developed a custom-designed, anaerobic, automated fraction collection system which we will describe in future work.

There is an inherent elution peak width at half-height associated with the chromatography column itself even for monodisperse molecules such as octane (Fig. 3).[30] Unlike reverse-phase chromatography columns, for SEC columns the instrumental band-broadening does not depend on elution time, but simply on the average size of the microgel particles used to pack the column. For example, the high-resolution columns used in previous studies were packed with porous 5-μm particles and had significantly narrower inherent band-broadening than less expensive columns packed with 10-μm particles of the same material. This column band-broadening is experimentally determined by using a known monodisperse sample and the same column used for study of the unknown samples. Using decane C10, we found $\Delta t_{1/2}$=0.25–.27 min while we found $\Delta t_{1/2}$=0.40 min for purified (99.9%,Strem Chemicals) buckyballs, C60 and C70.[30] From shape considerations, the latter standard seems more appropriate to compare to our spherical nanocluster samples. Any cluster elution peak whose width exceeds this value implies some polydispersity in the sample provided the sample does not have specific chemical interactions with the column. In fact, as shown in our previous work, reproduced in Fig. 4, our best metal cluster samples have $\Delta t_{1/2}$=0.30–0.35 min, slightly narrower than the best available C60, so there is some uncertainty ($\sim$0.05 min) as to which value best represents the inherent column broadening. Roughly speaking, the subpopulations of such a polydisperse sample with elution times lying outside $t_e \pm 0.4/2$ will be separated from the clusters represented by the signal at the apex of the peak and contain additional spectral information. The optical spectra obtained from the size-selected population eluting at the peak apex represents a single size to within the ±2 Å resolution for our column. This resolution of ±2 Å may be increased by increasing the total pore volume of the column by adding an identical column in series.

Typically, a 10-μl amount of $\sim$0.01 M nanocluster solution is injected from a crimp top vial using our Water's model 717 autoinjector and sampling loop. The total experimental time is 15 min, so complete size and optical analysis of 20 samples under automated control is quite feasible in a day. The retention time at the peak apex elution is measured by using the absorbance obtained

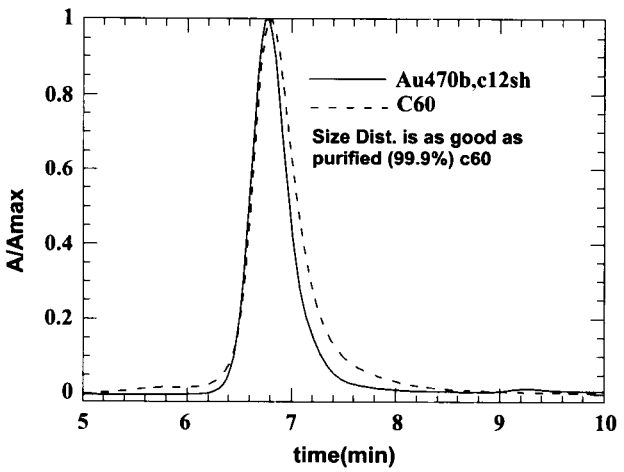


Fig. 4 Size distribution of Au clusters is as narrow as buckyballs. Normalized absorbance chromatogram from PDA detector (A/A_{max}) for C12SH-stabilized Au nanoclusters and 99.9% C60. HPLC conditions were a PL500 column using toluene as the mobile phase flowing at 1 mL/min. The C60 elution time has been shifted to coincide with that of the Au nanoclusters for ease of comparison. (From Ref. [30].)

from a chosen wavelength element of a 1024-channel, on-line, photodiode array (PDA), absorbance detector. The mobile-phase background absorbance, although not significant in the region of interest for most metal nanoclusters, is automatically removed. Organic chemicals (e.g., solvents, surfactants) that do not absorb at mobile-phase-transparent wavelengths, are detected by using an on-line differential refractometer (Water's model 410).

Transmission electron microscopy (TEM, HRTEM)

About 2 μl of the purified $\sim$0.01 M solution of nanoclusters was deposited on a holey carbon grid, under which a porous filter paper was placed to quickly wick the solvent away to minimize cluster pile up on the grid. For core sizes larger than $D_C \sim 2$ nm high-resolution TEM (HRTEM) was used to confirm the metal core size to an uncertainty of one lattice constant (±2 Å). However, in this size range, this measurement uncertainty is as large as the SEC resolution, even assuming a perfectly monodisperse sample with all the clusters in an identical focal plane on the holey carbon grid and identical nanocrystal orientation.

Nanocrystal orientation is very important for small crystals. For example, if a randomly oriented distribution of monodisperse icosohedral Au crystals is examined with $d\langle 111\rangle$=2.1 nm and $d\langle 100\rangle$=2.7 nm, one obtains the number average D_c(TEM)$\sim$2.4±.3 nm. but this is not a result of size dispersity. In addition, with the smallest clusters, especially Ag, there are e-beam cluster "melting"

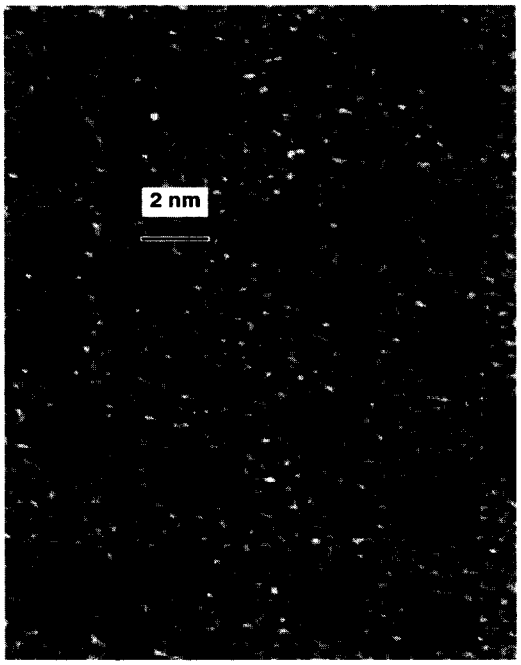


Fig. 5 HRTEM of a region of a holey carbon grid on which a drop of sample Au, $D=1.8(.2)$ nm was allowed to dry. (From Ref. [30].)

artifacts that make absolute size measurements difficult because the boundary with the grid becomes less sharp, in addition to causing some cluster fusion where their surfaces touch. For these reasons, size determination by high-resolution SEC, complemented by HRTEM, has major advantages. Fig. 5 shows an HRTEM of Au, $D_c=1.8$ nm particles, showing the atomic lattice fringes from this HPLC-purified sample, and illustrates the effects of crystallographic orientation and focal plane effects on the apparent cross-sectional diameter of this nearly monodisperse sample.

Absorbance spectra

The absorbance spectra of passivated metal nanoclusters was obtained by using an on-line photodiode array PDA with an adjustable bandwidth and wavelength range. For the purpose of the studies reported in this review, the bandwidth was either 2.4 or 4.8 nm, and a wavelength range of 290–795 nm was used. Complete absorbance spectra were collected every 2 sec during the chromatography. A Cary 2300 ultraviolet (UV)–visible spectrometer was also used to verify selected results from the PDA. Both spectrometers were calibrated by using a holmium filter.

Because a typical cluster elution linewidth is about 0.3–0.4 min, between 10 and 15 complete spectra were available as a function of elution time (i.e., size). Thus, in addition to the information concerning how the absorbance spectra changes with cluster size at the peak of the elution, we also can determine the spectral homogeneity of an elution peak. A spectrally pure or homogeneous elution peak corresponds to one in which there is no variation in the absorbance spectral shape within the peak. Size or shape polydispersity would lead to size-dependent optical absorption spectra within the elution peak.

Synthesis Optimization with SEC Feedback

Perfecting the methods of inorganic nanocluster synthesis requires rapid, quantitative feedback concerning final cluster size, shape, optical properties, and their relationship to the myriad synthetic variables, which can affect the average cluster size and size distribution. A good analogy is with advances in organic synthesis which have been facilitated greatly by the advent of modern analysis methods, such as Fourier-transform infrared (FTIR), gas chromatography/mass spectrometry (GC/MS), liquid chromatography (LC)/PDA, nuclear magnetic resonance (NMR), etc. Without such feedback, the synthetic chemist is very much in the dark with respect to whether changes in synthetic protocol are making things better or worse. Another example comes from the application of liquid chromatography/mass spectrometry (LC/MS) and capillary gel electrophoresis approaches to DNA and protein analysis which, along with polymerase chain reaction (PCR) amplification, essentially enabled the rapid sequencing of the human genome. Unfortunately, carrying such analytic methods directly over into the field of inorganic nanocluster synthesis is not always straightforward. Perhaps, this is the reason for the limited number of nanomaterials that humans know how to synthesize in a predictable manner—the best examples being transition metal nanoclusters as pioneered by my group, and II–VI semiconductors (mainly CdSe) as pioneered by Murray, Bawendi, Alivisatos, and Brus.[6,38–40] These materials were explicitly chosen because they have clearly identifiable, size-dependent visible colors (absorbance) that allows rapid feedback regarding size and monodispersity. For example, in the case of CdSe, size changes as small as 1–2 Å result in clear differences in the absorbance edge. In the case of nanosize Au or Ag, one can perceive, even with the naked eye, color changes corresponding to ~2–4 Å in size, and using an absorbance spectrometer, 1–2 Å size differences.

RESULTS AND DISCUSSION

Our recognition of the dearth of rapid analysis techniques for inorganic nanomaterials led us to try to adapt some modern analytical methods to understand size/optical

property relationships in both semiconductors and nanometals. These techniques include in situ elemental analysis of purified cluster solutions using X-ray fluorescence (XRF), which is very useful for composite nanoclusters, and size-exclusion liquid chromatography, which has recently become our most useful tool, supplanting even TEM in both size precision and utility. We now illustrate the power of these new analysis methods in improving nanocluster synthesis with several examples from our previous work.

Nanocluster/Ligand Binding Studies

We have recognized for many years that at least four important factors control the final size of nanoclusters synthesized in inverse micelles:

1. the micelle size,
2. the binding strength of the surfactant micelle used,
3. the amount of precursor salt used in the synthesis,
4. the strength of the reducing agent used.

Variables 3) and 4) influence the final cluster size even in the absence of micelles as, e.g., in metathesis methods using metal–organic precursors in the presence of coordinating ligands. Perhaps, not surprisingly, increasing the concentration of metal salt precursor increases the final cluster size, while increasing the rate of reduction by using a stronger reducing agent decreases the cluster size. Also, increasing the strength of the binding of the micelle-forming surfactant decreases the final cluster size if all other variables remain fixed. However, it is very difficult to predict, a priori, which surfactants will bind most strongly to which metals, nor how much change in final size will be achieved by changing the precursor concentration or reducing strength.

By measuring the total amount of clusters eluting from a given column/mobile phase using measurement of the elution area at a given wavelength with and without the column, one can study the competition for the organic ligand between the high surface area column and the high surface area nanoclusters. It is found that, for Au nanoclusters, thiol exchange does not occur on the time scale of the chromatography experiment ($\sim$15 min) even in the presence of an overwhelming amount of a different CkSH in the mobile phase [i.e., the clusters elute at the time expected for the chain length (thickness) of the alkyl thiol], while for Ag nanoclusters, complete exchange occurs under these conditions. Thus, Ag nanoclusters bind alkyl thiols significantly more weakly than Au. Pt, Pd, and Rh nanoclusters do not exchange thiols during the chromatography like Ag, but still have slightly more weak binding than in the case of Au clusters of the same size, as indicated by fewer (<100%) eluting clusters.[30]

One can learn about ligand metal cluster binding, so critical to catalysis, by examining the relative amount of clusters eluting from a column with a mobile phase of pure toluene. In this case, the column attempts to separate the ligand from the cluster. Failure to do so results in complete elution of the clusters, as judged from ratio of the area under the elution peak compared to the no-column case. In Fig. 6, one observes that the relative areas decrease with k for k=6, 10, and 14, indicating shorter chain thiols bind less strongly than longer ones.[30]

These types of subtle binding effects are nearly impossible to study by other methods. In fact, the organic ligand cannot be seen in TEM so its length can only be indirectly inferred from the intercluster gap. An interesting prediction from our SEC ligand/cluster binding studies is that, under vacuum as in an electron microscope, only longer chain thiols will be able to retain their ligands, preventing cluster coalescence and allowing the formation of hexagonal or cubic quantum dot arrays (QDAs). This is indeed the case as we have shown in previous work on QDAs.[41,42] Another prediction based upon our SEC work is that Ag nanoclusters will more readily desorb even long chain alkyl thiol ligands under vacuum and so, the longevity of QDAs of Ag under TEM is significantly decreased relative to Au. Also, electron beam melting effects are more severe in QDAs of Ag than of Au although the melting temperatures of the two materials are very similar. As expected, QDAs of alkyl thiol ligated Pt, Pd, and Rh are more robust than Ag under vacuum.

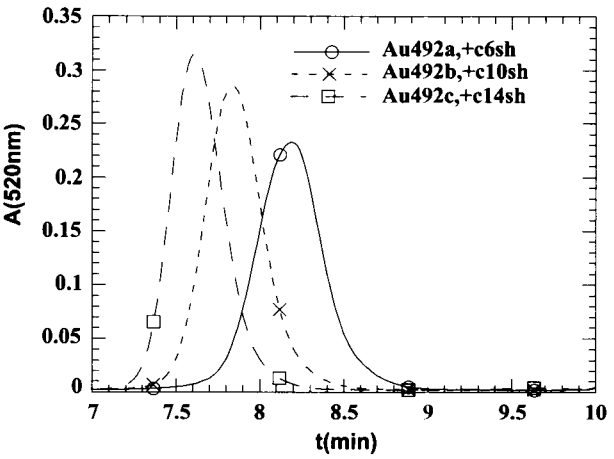


Fig. 6 Effect of capping agent size on the hydrodynamic diameter of Au nanoclusters. HPLC chromatogram of Au nanoclusters with three alkyl thiols, CkSH, k=6, 10, 14, added after reduction. The absorbance at 520 nm [A(520nm)] from the PDA vs. elution time is shown. The SEC column used was a Polymer Labs PL1000 and the mobile phase was toluene at 1 mL/min flow rate. (From Ref. [30].)

A somewhat surprising observation common to all metal Qdots, which we have studied, is that monodentite ligands such as alkyl thiols, alkyl sulfides, and primary amines bind increasingly less effectively as the size of the nanocluster increases. In fact, for $D_c > 10$ nm, one finds that displacement of a polydentite ligand such as a nonionic alkylated polyethylene oxide by a more strongly binding alkyl thiol will result in complete irreversible aggregation of a Au or Ag nanocluster solution! Thus, one needs to develop more advanced polydentite ligands for larger nanoclusters.

Intuition often fails when choosing an appropriate multidentite ligand for a given metal Qdot. For example, we have found that alkyl chain disulfides aggregate these larger clusters, which surprised us greatly. This observation probably indicates that the hypothesis that alkyl thiolates form disulfides on nanocrystalline gold is wrong. Fortunately, nonionic block copolymers, such as were employed in the first synthesis of Co via high-T decomposition,[5] are obvious choices and, if selected appropriately, they stabilize a very broad range of transition and base metal Qdots extremely effectively. However, they do not allow efficient SEC analysis of the stabilized large nanoclusters as most columns are too effective at pulling them off the cluster surface, and such bulky stabilizers may prevent deposition of additional atoms to form core–shell particles, so they are not a panacea. Thus we are still searching for the most effective steric ligands for SEC analysis of large, $D_c > 10$ nm nanoclusters.

SEC Studies of Synthetic Variables

Traditionally, TEM or X-ray diffraction (XRD) was used to infer the nanocluster particle size. In the case of the former, between a few dozen particles to a few hundred are measured from a "chosen" area of the grid and tediously counted. Many "representative" areas of the grid must be chosen by the microscopist, requiring significant expertise and time. To obtain a detailed analysis of even a single sample in a day's time requires a dedicated graduate student or postdoctorate individual—highly motivated by the possibility of a permanent job! Considering that a skilled synthetic inorganic chemist can easily make a dozen samples in the same period of time, one can see where the roadblock to synthetic progress occurs.

Other tradition surface science methods of analysis often prove inadequate as well. For example, in XRD characterization of size, the first maximum in the low-angle diffraction pattern can be used to obtain the average interparticle spacing (assuming either hexagonal or cubic packing). Unfortunately, the interparticle gap depends on the organic ligand and its degree of interdigitization, so this method of core size determination is not ideal, although it is applicable to the smallest clusters with $D_c < 2$ nm, and has been used with good precision for very monodisperse clusters.[43]

As we have extensive experience with small-angle neutron scattering (SANS), X-ray scattering (SAXS), and dynamic light scattering (DLS), I initially believed these methods would prove very useful for size and size dispersion analysis.[44] However, in the case of metal Qdots and DLS using visible light, the very strong absorbance and concomitant weak scattering requires very low nanocluster concentrations so that the signal-to-noise (S/N) ratio is exceptionally low for clusters with $D_c < 4$ nm. This technical difficulty has precluded our routine use of this method. In the case of SANS, the scattering length density difference between solvent and typical metal Qdots is very low, again giving low S/N and long data collection times. SAXS is the best approach with respect to S/N from metallic Qdots, but typically requires an intense tunable (synchrotron) X-ray source, to avoid large (e.g., 24 hr) data collection times.[44] Thus until we had developed SEC for nanoclusters, our synthetic feedback was poor.

To illustrate the speed, precision, and high information content of SEC size/optical property analysis, we now review some selected cases of using SEC to determine the effect of variation of synthetic parameters on metal Qdot synthesis. We refer the reader to the results of a recently submitted manuscript for further examples.[31] First, we fix the solvent, metal precursor concentration, reductant,

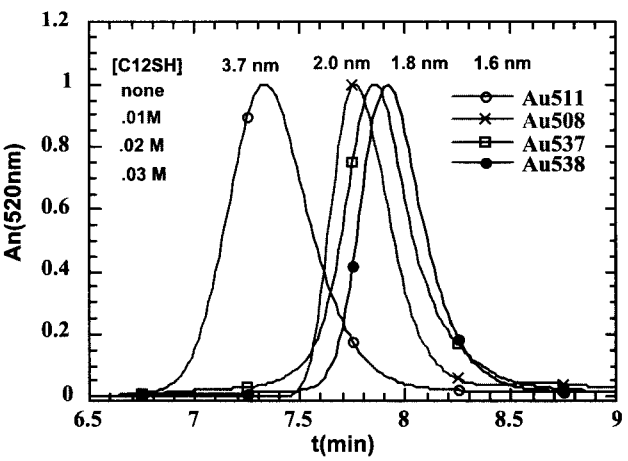


Fig. 7 Effect of C12SH concentration on the final nanocluster size and size distribution of Au nanoclusters made by $NaBH_4$ reduction of a two-component TOAB/toluene inverse micelle system with [Au]=0.01 M. The column is a PL1000 type and the mobile phase is toluene with C12SH added at 0.01 M. The SEC core sizes assume a core-size-independent 2.4-nm total organic shell, and are indicated above each curve. (From Ref. [31].) (*View this art in color at www.dekker.com.*)

and inverse micelle surfactant, and examine the effect of alkyl thiol concentration on the final cluster size. Fig. 7 shows the SEC chromatograms from Au nanoclusters synthesized with various amounts of added C12SH from 0.0 to 0.03 M. As expected, having more of the tightly binding alkyl thiol restricts the growth more strongly. The linewidths observed (size dispersions) only vary slightly once the thiol concentration exceeds 0.01 M. (The Au/C12SH ratio is 1:1 for this case.) In fact, once the Au/CkSH ratio exceeds 1.0, the samples are nearly monodisperse. It is noteworthy that although TEM easily gives the size of the largest cluster made with no alkyl thiol present, Au #511, as $D_c \sim 3.8$ nm, the other three samples all appear to be around $D_c \sim 1.8 \pm .2$ nm using HRTEM because identification of the cluster/substrate boundary combined with orientation effects obscure the true size differences, which SEC clearly demonstrates. (2 Å is one lattice fringe and thus is the inherent measurement uncertainty for a perfectly monodisperse sample.) Also noteworthy is that this information was obtained in 15 min/chromatogram by using only 10 μl of a 0.01 M Au cluster sample! On this particular day, the effect of over 12 different synthetic variables were determined by SEC.

We obtain the complete absorbance spectra from our PDA array throughout the peak elution (every 2 sec, which corresponds to a size discrimination of only 1–2 Å). The spectra obtained at the peak apex represent the absorbance of the majority of the nanoclusters. The spectra normalized at a common value, 500 nm, are shown in Fig. 8 for each sample. One observes the characteristic blue shift of the plasmon, which we first identified in Ref. [13] and confirmed using SEC in a recent paper.[36] As the size decreases from ~2.0 to 1.6 nm, additional

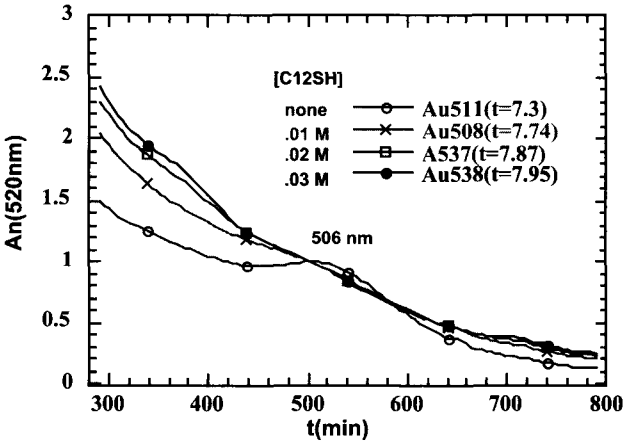


Fig. 8 Absorbance at the elution peak of each of the chromatograms of Fig. 7, illustrating the blue shift and plasmon broadening with decreasing size. All the clusters are in toluene, which is also the mobile phase. (From Ref. [31].) (*View this art in color at www.dekker.com.*)

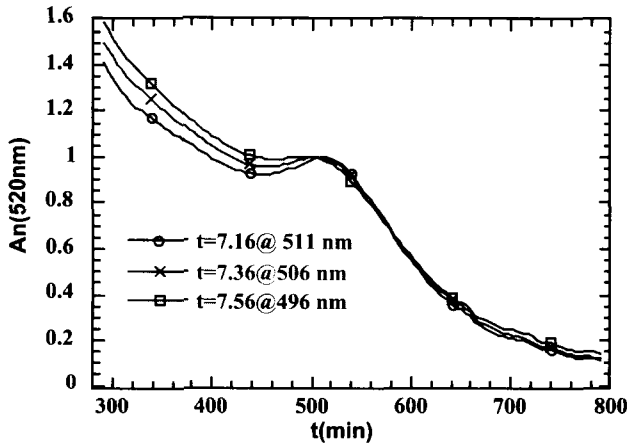


Fig. 9 Elution peak spectral homogeneity of the sample Au, $D_c = 3.8$ nm, with the broadest size distribution (linewidth) shows the population to be mildly inhomogeneous in size, but with only a spread of plasmon positions from 496 to 511 nm, a size dispersion of about ±0.5 nm. (Previously unpublished work.) (*View this art in color at www.dekker.com.*)

nonclassical features appear in the absorbance as a result of the discrete density of states of a gold cluster. For example, $D_c = 1.6$ nm corresponds very closely to the magic, two fully closed shells icosohedra with $N = 55$ having $D_c = 1.5$–1.6 nm first synthesized and identified by Schmid.[7] Recently, Whetten has also identified this magic size using MALDI TOF masss spectroscopy.[45,46] The features of this smallest ~9 kDa cluster fraction correspond closely to the spectra of Au #538 shown in Fig. 8. The enhancement of both the UV and near-infrared (NIR) absorbance is characteristic of the quantum confinement effects in ultrasmall metallic Qdots[a] and corresponds to the absorbance features which are observed in semiconductor Qdots.[6,47]

Both Au and Ag possess strongly size-dependent optical absorbance in the quantum confinement regime, $D_c < 5$ nm, which allows us to assay the size homogeneity of the elution peaks of Fig. 7 by examining the "spectral homogeneity" of the peak. Measuring the uniformity of the spectral shapes of an eluting component is a commonly used approach to look for possible coelution of inhomogeneous species even with simple chemicals.[37] We have extended its use to Au, Ag, and Cu nanoclusters. In the case of metallic Qdots of Au or. Ag, spectral

[a]Quantum size effects are a result of the discrete nature of the electronic levels as a cluster is built up from atoms. The level splitting (ΔE is roughly given by $\Delta E \sim E_f/N$, where E_f is the Fermi level of the metal and N is the number of atoms in the cluster. If $E_f \sim 5$ eV, then ($\Delta E \sim 50$ meV for a 100-atom cluster. At room temperature, one might see these effects for Au at $N \sim 400$ or $d \sim 3$ nm.

inhomogeneity could arise from either size or shape effects. In fact, we find that only the largest component, Au D_c=3.8 nm, made without C12SH present during the synthesis, possesses any significant spectral inhomogeneity, as demonstrated by the spectra obtained at the peak apex time, t=7.36 min, and at the half-width at half maximum positions, t=7.16 and t=7.56 min (Fig. 9). The total variation of the plasmon position even for this sample is incredibly small, as the bandwidth of the detector was set to ±4.8 nm for these experiments. However, it is certainly real and corresponds to size inhomogeneity of ~±0.5 nm. Samples made with alkyl thiol present during the reduction have elution linewidths and spectral homogeneities consistent with complete monodispersity—a tribute to both the inverse micelle method and the role of alkyl thiols as etchants to narrow the size distribution.[31,45]

What is the role of inverse micelle type and solvent on the final cluster size? Both of these issues were the subject of a recent publication.[31] In that work, it was demonstrated that even in the presence of a strongly binding thiol, the size and/or strength of the micellar binding to the encapsulated precursor salt had a significant effect on the final size and size distribution. Fig. 10 shows the cluster size distribution obtained without deconvolution of the instrumental linewidth for nonionic, three types of cationic surfactant, and an anionic surfactant. In each case, the solvent was held constant, toluene, as was the reductant, $LiAlH_4$.

As was noted in the original reference, the structure and binding properties of the inverse micelle used is a competitive factor with the binding of the alkyl thiol surfactant. For example, the presence of the significantly weaker binding tetrahexylammonium chloride (THAC) results in much larger clusters than do bromide head group cationics didodecyldimethylammonium bromide (DDAB) or tetraoctylammonium bromide (TOAB). Didodecyldimethylammonium bromide has a smaller micelle interior volume and one could rationalize the smaller final cluster size compared to TOAB by this factor.

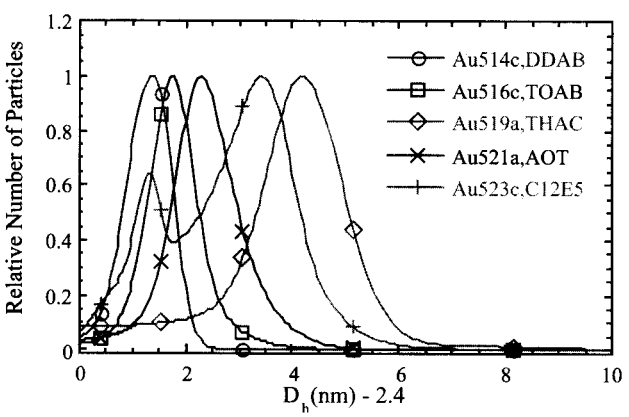


Fig. 10 Plot of the Au core size $D_c=D_h-2.4$ obtained by assuming the previously measured total shell thickness of 2.4 nm, vs. normalized detector response at 520 nm (relative number of particles). The surfactants used in the synthesis are indicated. (From Ref. [31].) (*View this art in color at www.dekker.com.*)

A most surprising result was the formation of a bimodal population distribution from inverse micelles of the nonionic surfactant C12E5 in toluene, whereas only a narrow dispersion, monomodal distribution is produced when alkanes are used as the oil continuous phase, as has been reported previously.[35,36] This illustrates that the packing of the surfactant molecules into a droplet-like interface, and the lability of this interface is strongly affected by the interaction of the surfactant alkyl chains with the oil used.

Finally, one might wonder about the role of precursor salt concentration on the final size and size distribution in the inverse micelle synthesis. To study this we used a single cationic surfactant two component micelle system (TOPB/tol), where TOPB=tetraoctylphosphosium bromide and a fixed [metal]/[surfactant] ratio. As the [HAuCl4] was varied from 0.01 to 0.1 M, no discernible change in size from D_c~1.9 nm was observed, nor did the size distribution alter as monitored by elution peak linewidth. This is in strong contrast to metastatic reduction reactions of organometallics in coordinating organic solvents, where an increase in [precursor metal–organic] leads to systematically greater final sizes with concomitant increases in size dispersion. For example, in the case of Fe(II) reduction in a coordinating solvent, an increase in size from D_c~2.0 to 14 nm occurs over the same concentration range. This illustrates the fundamentally different mechanism operating in the inverse micelle case.

Optical Properties of Au, Ag, and Au/Ag Core–Shell Nanocrystals

In the proceeding section, we touched on the very interesting optical properties of Au, which exhibits a nonclassical (i.e., not explained by the Mie theory of electromagnetic scattering from metal colloids), blue shift and damping of the plasmon absorbance peak with decreasing size.

In Fig. 11, we show the absorbance chromatograms of several samples of Au nanocluster samples all stabilized with dodecanthiol to ensure a nearly constant shell thickness and interface environment. The peak heights have all been normalized to allow easy comparison of the elution shapes and widths. The elution peaks have *not* been deconvoluted with the instrumental band broadening. The size, as determined from the peak apex elution time, is given in Table 2. A constant shell thickness of

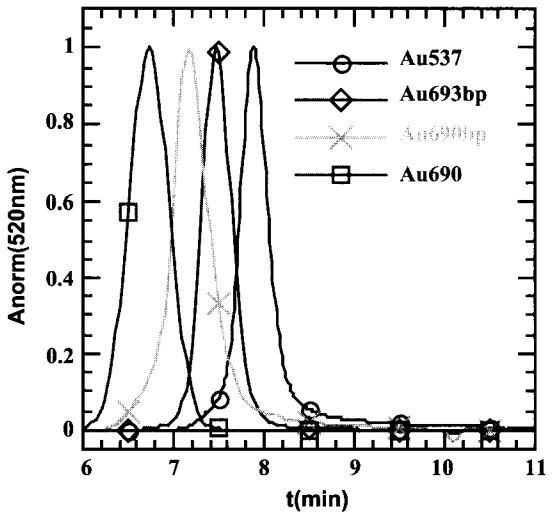


Fig. 11 Normalized absorbance chromatograms of several differently sized Au nanocluster samples. (From Ref. [36].) (*View this art in color at www.dekker.com.*)

2.4±0.1 nm, as determined previously,[30] was used for all samples to obtain the metal core size.

The corresponding optical spectra for each sample at the peak of its elution are shown in Fig. 12. The spectra are normalized so that absorbance=1.0 at 500 nm in each case. From this figure, we obtain the size-dependent blue shift and broadening of the optical plasmon as a function of size. This information is contained in Table 2.

The data from Fig. 12 and Table 2 demonstrate that, as the size decreases into the quantum size regime, a very rapid broadening and blue shift of the absorbance occurs between D_c=3.0 nm and D_c=1.3 nm. The classical Mie theory assumes no size dependence to the complex dielectric constant $\varepsilon(\omega)$ and so predicts that the optical plasmon for spherical Au clusters will blue shift with decreasing size until a size of $D_c \sim 10$ nm is reached, at which point no further changes from the predicted peak position of 530 nm (in toluene) or half-width should occur. Thus all the spectra shown in Fig. 12 are nonclassical in character, exhibiting significant blue shifts and broadening compared to the Mie prediction. The effects of cluster size are usually accounted for by assuming that $\varepsilon(\omega)$ is size-dependent. This allows classical effects, such as increased electron scattering at the metal–dielectric interface, to be empirically incorporated into the theory, as well as quantum size effects (QSE) and the effect of chemical interface damping. Although all these effects are predicted to increase the linewidth, just as what has been observed in the present experiments, as well as the majority of experiments on embedded clusters, a prediction of the direction of the peak shift is much more problematic.[2]

The absorbance spectra in Fig. 12 have been normalized at a common wavelength, 500 nm, to allow easy comparison of the peak shapes. (See Table 1 for sample synthesis conditions and other details. Note that the sample number has nothing to do with the number of atoms in a cluster, which can only be approximately estimated even with TOF-MS). Both the UV and, to a lesser extent, the NIR absorbance are enhanced as the cluster size decreases. We have also noted the development of new features in the absorbance profile of the smallest D_c=1.7 and D_c=1.3 nm clusters.[31,36] More recent work by others has confirmed this trend, which is attributed to the appearance of molecule-like density of states in small metal clusters.[45,46]

One does not need separate, completely monodisperse samples to determine the optical absorbance behavior of metal nanoclusters as a function of size. because the SEC process will separate clusters with very small differences in size $\sim\pm 2$ Å and the PDA detector will provide the spectra corresponding to these sizes. The information obtained on samples exhibiting size polydispersity is very consistent with that obtained from the peak apex spectra of the individual samples of Table 2, demonstrating the consistency of the SEC size analysis. For example, in Table 2, we show that Au #693bp has a peak retention time of 7.5 min corresponding to an average size of 2.9 nm and a plasmon peak at 501 nm. This is within the experimental uncertainty of the data of Fig. 12 corresponding to a size of 2.8 nm and having a plasmon peak

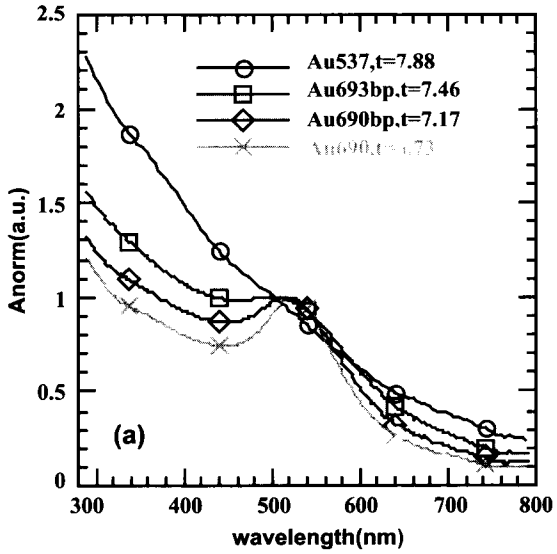


Fig. 12 Normalized peak elution spectra from Fig. 11 demonstrate the effect of decreasing Au nanoclusters size (increasing elution time) on the optical absorbance in the visible. (From Ref. [36].) (*View this art in color at www.dekker.com.*)

Table 1 Ag and Au nanocluster synthesis

Sample name	Metal salt	Surfactant	Solvent	Reductant	Passivator
Au511p	HAuCl4 (0.01 M)	TOAB (0.06 M)	Toluene	$NaBH_4$ (0.04 M)	C12SH
Au514cp	HAuCl4 (0.01 M)	DDAB (0.1 M)	Toluene	$LiAlH_4$ (0.04 M)	C12SH
Au521a	HAuCl4 (0.01 M)	AOT (0.2 M)	Toluene	$LiAlH_4$ (0.04 M)	C12SH
Au537	HAuCl4 (0.01 M)	TOAB (0.06 M) C12SH (0.03 M)	Toluene	$NaBH_4$ (0.1 M)	N.A.
Au567p	HAuCl4 (0.01 M)	TOAB (0.06 M) C16SH (0.03 M)	Toluene	$LiAlH_4$ (0.04 M)	N.A.
Au687bp	HAuCl4 (0.01 M)	C12E5 (0.2 M)	Octane	LS-SelBH (0.04 M)	C12SH
Au690	HAuCl4 (0.01 M)	C12E5 (0.2 M)	Hexadecane	SuperH (0.08 M)	N.A.
Au690bp	HAuCl4 (0.01 M)	C12E5 (0.2 M)	Hexadecane	SuperH (0.08 M)	C12SH
Au693bp	HAuCl4 (0.01 M)	C12E5 (0.2 M)	Hexadecane	LSelAlH (0.08 M)	C12SH
Ag114	AgNO3 (0.01 M)	TOAB (0.12 M)	Toluene	$NaBH_4$ (0.02 M)	N.A.
Ag205	AgNO3 (0.01 M)	TOAC (0.12 M)	Toluene	$LiAlH_4$ (0.02 M)	N.A.
Ag213	AgNO3 (0.01 M)	TOPB (0.12 M)	Toluene	$NaBH_4$ (0.02 M)	N.A.

Alkanes of chain length k are abbreviated Ck, alkane thiols CkSH. Surfactant abbreviations are: AOT = ***bis***(2-ethylhexyl) sulfosuccinate sodium salt, C12E5 = penta-ethyleneglycol-mono-*n*-dodecyl ether, DDAB = didodecyldimethylammonium bromide, TOAB = tetraoctylammonium bromide, TOAC = tetraoctylammonium chloride; reductants prepared in anhydrous tetrahydrofuran (THF) as 1–2 M solutions, SuperH = $LiB(C_2H_5)_3H$, LS-SelBH = $LiB(CH(CH_3)CH(CH_3)_2)_3H$; LSelAlH = $LiAl[OC(CH_3)_3]_3H$, final concentrations are indicated in table.

at 496 nm because the spectra bandwidth was set at 2.4 nm. The size-selected cluster population of Au #511b, which elutes near the half-maximum at t=7.14 min, corresponds to a SEC size of 4.2 nm and has identical spectra to that at the peak apex of Au #690bp.

In a recently submitted paper, we demonstrate a synthesis based upon the seminal work of G. Schmid,[7] which illustrates the selection and optical characterization abilities of HRSEC for strongly polydisperse samples.[31] In that paper, we illustrated the effects of the emergence of a discrete, "molecule-like" density of states in Au "molecules" specifically, Au(N=13), the smallest closed shell atomic configuration. Such molecular species had been previously shown by us[48] and subsequently by others[45,46] to have both visible and NIR photoluminescence (PL), which is a startling result! This relatively

Table 2 Size-dependent optical properties of Au and Ag nanoclusters

Sample name	t_e (min)	D_{core} (nm, SEC)	D_{core} (nm, TEM)	λ_p (nm)	$\Delta\lambda_{1/2}$ (nm)
Au511p	7.24±.21	3.60	3.5	506±2.4	110
Au514cp	8.02±.19	1.33	1.5	none	
Au521a	7.68±.21	2.21	2.4	none	
Au537	7.88±.19	1.67	1.8	none	
Au567p	7.80±.05	1.40	1.5	none	
Au687bp	7.50±.02	3.10	3.0	496±2.4	120
Au690	6.73±.30	6.0	6.2	523±2.4	69
Au690bp	7.17±.20	4.0	4.2	513±2.4	86
Au693bp	7.47±.26	2.90	3.0	501±2.4	120
Ag114	7.93±.17	1.55	1.8	486±4.8	124
Ag205	7.24±.22	3.70	4.0	440±4.8	60
Ag213	6.3±.25	7.6	7.5	423±4.8	39
	7.73±.17	2.4	2.5	474±4.8	91
Ag213	6.7	6.2	Not measured	423±4.8	48
	7.0	4.7		428±4.8	69
	7.2	3.9		443±4.8	76
	7.4	3.1		457±4.8	79
	7.6	2.5		467±4.8	86

$\Delta\lambda_{1/2}$ (nm)=Half-width at half maximum determined on long wavelength side of asymmetrical plasmon peak.

strong PL ($\sim 10^{-4}$ Q.E.) can come from both s-to-d interband transitions (visible PL), or intra-d-band transitions (NIR and IR PL). Such unexpected optoelectronic properties illustrate the reason why very small nanoclusters are so interesting.

Ag nanoclusters

We next examine the size dependence of the optical absorbance of Ag nanoclusters grown in inverse micelles and stabilized with the identical alkyl thiol (C12SH) used in our Au cluster studies. A strong motivation for this complementary study is the important difference in the optics of Au and Ag clusters arising from the different energies of the interband (filled d shell to conduction sp band) transition onsets in the two metals. The close proximity of the interband transition to the plasmon energy in Au (it lies to the red of the plasmon energy) is the reason why energy damping (peak broadening) is so much greater for Au clusters than for Ag, where the interband transition energy onset lies significantly to the blue of the conduction band plasmon. This makes a determination of the size dependence of the linewidth for small Au clusters difficult (i.e., the peak simply disappears). It also leads to a pronounced asymmetry to the peak shape.

There has been a matter of significant controversy regarding the direction of the plasmon absorbance energy shift with cluster size in the quantum size region. Depending on details and relative weights assigned to the various theoretical factors, one can "predict" either a blue or a red shift with decreasing size. We had demonstrated a blue shift in the case of monodisperse, spherical nanoclusters of gold, and this is the general consensus based on dozens of papers in this field. However, the case of Ag is not as well established, and both red and blue shifts have been reported. We were considerably surprised when SEC and on-line optical characterization of our inverse micelle prepared Ag nanoclusters showed a clear red shift with decreasing size. The magnitude of this shift was truly astounding compared to gold. In the smallest clusters we studied, the damping of the very sharp and relatively symmetrical absorbance plasmon of Ag completely disappears—the first such observation in nanosize Ag. Fig. 13 shows the effect of size decreases on the position of the optical absorbance of Ag nanoclusters.

The main feature of the optical absorbance of nearly free-electron metals such as Cu, Au, or Ag is a condensation of the conduction electron oscillator strength into a narrow frequency band of plasmon polariton excitations, which are caused by the particle surface—the plasmon resonance. If the dielectric function for a metal cluster does not depend on its size, little or no variation in the position of energy maximum of the

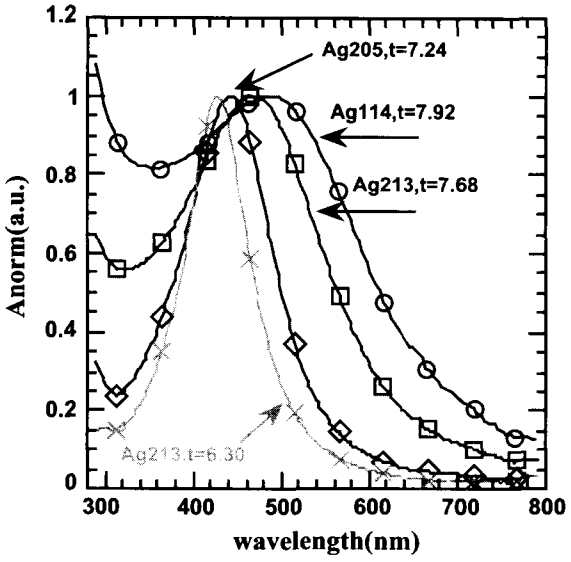


Fig. 13 SEC separation and on-line characterization of Ag nanocluster solutions in toluene show the strong red-shift and extreme damping of the plasmon with decreasing size. See Table 1 for synthesis details and other information. (From Ref. [36].) (*View this art in color at www.dekker.com.*)

plasmon absorbance profile is predicted to occur below a size of 10 nm. So, for particles smaller than this size, as Kreibig and Genzel have pointed out,[49] the absorbance band shape only changes with size if intrinsic particle effects are important. It has been established by numerous experiments that, for sizes less than ~ 5 nm, $n(\omega,R) \sim 1/R$.[2] This is, of course, just the surface-to-volume ratio.

Both classical effects [such as electron scattering at the metal/dielectric interface and quantum size effects (QSE)] become important at sizes below ~ 5 nm. The latter effects, in particular, should become important when the level spacing near the Fermi level exceeds the available thermal energy. This will occur at about 400 atoms of Au, or a size of about 2.5 nm, assuming the nanoclusters to have the density of the bulk material.[a] It turns out that both classical effects, such as the reduced free path of the electron and QSE effects, both predict a $1/R$ dependence to the resonance linewidth, so an observation of a $1/R$ behavior to the resonant linewidth does not, by itself, demonstrate a QSE. However, the emerging weak features in the spectra of our smallest clusters probably do indicate the onset of discrete bands, which is a true QSE. It turns out that both classical effects, such as the reduced free path of the electron and QSE effects, both predict a $1/R$ dependence to the resonance linewidth, so an observation of a $1/R$ behavior to the resonant linewidth does not, by itself, demonstrate a QSE. However, the emerging weak features in the spectra of our smallest clusters probably do indicate the onset of discrete bands, which is a true QSE.

A further complication in the interpretation of experimental observations in either classical (electron scattering) or quantum confinement (continuous bands becoming discrete) behavior occurs in the case of metals such as Au and Cu, in which the onset frequency of the interband transitions from d-type orbitals to the sp-type conduction band is close to the plasmon energy. In this case, there are major changes in the resonance energy just because of the importance of these transitions. For example, in the case of both Cu and Ag, the free-electron (Druid model) Mie theory predicts a resonance of $\sim$9.2 eV in a vacuum ($\varepsilon_m = 1$); however, the contribution of the 4d core electrons to the susceptibility shifts this to the experimental value of $\sim$3.8 eV for Ag but $\sim$2 eV for Cu! It may be that the different directions of the plasmon energy shift we have established for Au and Ag with the same passivating layer, embedded in an identical media, toluene, are a result of the relatively greater importance of the interband transitions in Au compared to Ag. For a more extensive discussion of our findings for nanosize Au and Ag, we refer the reader to Ref. [36] and the monograph by Vollmer and Kreibig,[2] which is the definitive modern work on this subject.

The Ag samples whose spectra were shown in Fig. 13 are all coated with an identical ligand, C12SH, which gives keeps the clusters from aggregating even when deposited onto a holey carbon grid and inserted into the high vacuum of an electron microscope. Fig. 14 shows a TEM of Ag #205 illustrating the ready formation of arrays which we have termed quantum dot arrays—the next subject of this review.

We summarize our findings concerning the size-dependent optical properties of Au and Ag as follows. For both Au and Ag clusters in the size range $D_c = 8$ to $d = 1.5$ nm, the plasmon linewidth broadens following a $1/R$ linewidth size dependence, whose slope is greatest for Au. The peak asymmetry in the plasmon band shape is greatest for Au and increases with decreasing size for both Au and Ag clusters. The plasmon peak energy blue shifts with decreasing size for Au clusters, while in the case of Ag nanoclusters a red shift is observed.

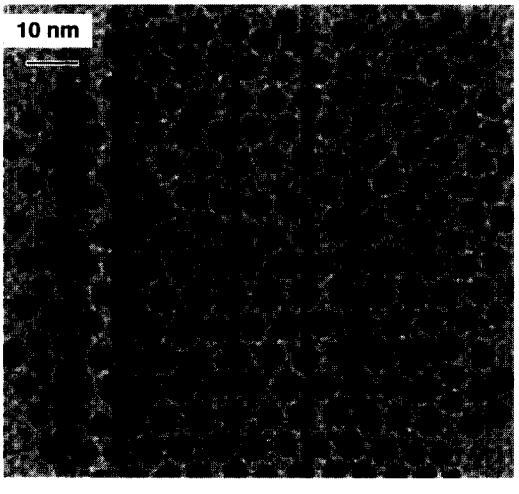


Fig. 14 TEM of Ag #205, $D_c = 4.0 \pm 0.5$ nm nanoclusters with a plasmon peak at 440 nm in toluene.

Core–Shell Particles and Alloys of Au and Ag

There are many scientific reasons to investigate core–shell or alloy nanoclusters. For example, even monolayer shell coverages can shift the Fermi level, E_f, via e-donation/acceptance by huge amounts (e.g., $\sim$20–30% donation/atom blue shifts E_f by $\sim$1 eV!). In the case of metals with dissimilar electron affinities, interface structures may form, which are similar to doping-induced depletion zones in bulk semiconductor materials. This effect may be a basis for metal rectification or spatial charge separation and thus of importance to nanocluster photocatalysis.[50,51] In view of the very interesting physics of such core–shell structures and the possibility of "reconstruction" occurring as the result of a heteroatomic deposition process, this area of nanometals should be very fecund and deserving of future investigation.

The coinage metals Au and Ag are a nearly an ideal system to study the effects of size and composition on the optical properties of nanoparticles because they have identical covalent radii, have simple fcc cubic lattice structures, and are miscible in all proportions. (However, this also means that diffraction-based methods such as TEM are not very useful for determining their structures.) One may pose a myriad of questions—some of which may even be theoretically treated—albeit, within the confines of classical electromagnetic Mie theory. For example, for fixed composition and size, how does the order of deposition Au/Ag vs. Ag/Au (core–shell) affect the optical spectra? For equivalent atomic compositions and size, does a random nanocluster alloy of Au and Ag differ optically from the corresponding core–shell cluster? For a fixed total size, how does composition affect the optical absorbance?

First, a brief remark on the synthesis of core–shell vs. alloy nanometals. To make alloy-type materials, a chosen ratio of precursor salts (typically $NaAuCl_4$ and $AgNO_3$) are dissolved in a suitable inverse micelle solution. A reducing agent is then added to effect coreduction. The agent chosen should approximately have the same reduction kinetics to avoid prenucleation of one material over the other, which would result in a core–shell type structure. By using an extremely aggressive reduction such as $LiAlH_4$, both metals undergo instantaneous reduction provided aggressive mixing is utilized.

To make core–shell structures, a simple variation on our newly developed solution epitaxial growth, which will be reported in a future work,[52] is employed. Specifically, a given seed nanocrystal, say $D_c = 1.8$ nm Au, is purified via SEC or precipitation methods, and a metal organic precursor solution is slowly coinjected along with the reductant, typically $NaBH_4$, via a syringe pump. A slow injection rate, typically 1–2 mL/hr is used, to ensure hetero- vs. homogeneous nucleation of the growth solution. A key aspect is the choice of a suitable surface active agent as well as the organometallic atom source. In general, one wishes to avoid the use of micelle-forming surfactants because these may act as nucleation centers for homoatomic clusters from the feedstock. For Au/Ag and Ag/Ag core–shell nanocrystals, small amounts of long-chain alkyl thiols work quite well and allow facile SEC analysis to be performed.

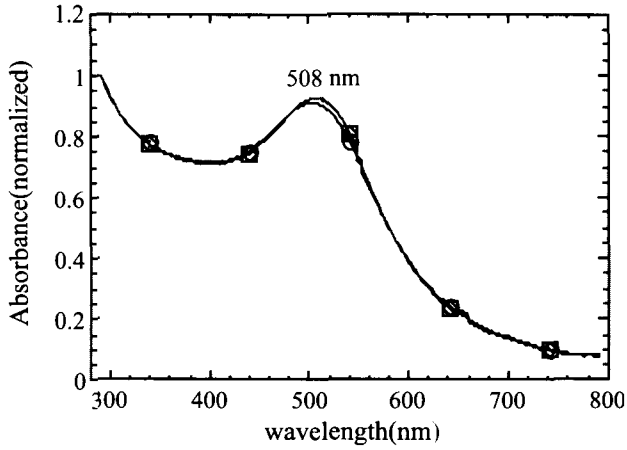


Fig. 16 The peak apex absorbance spectra of a $D_c = 2.4$ nm 1:1 Au/Ag alloy. (Unpublished data.) (*View this art in color at www.dekker.com.*)

Effect of Alloying on Ligand Binding

Size exclusion chromatography can be used to determine the relative binding of ligands of a given size and metal type as mentioned earlier in this review. We demonstrated that no thiol exchange occurs for CkSH, $k = 6$, 10, 14, in the case of Au nanoclusters, when an overwhelming amount of C12SH is in the mobile phase. In contrast, Ag nanoclusters with the identical thiols on their surface completely exchanged with the C12SH in the mobile phase during the chromatography as evidenced by a common elution time in all cases, which corresponds to a shell of C12SH (the core size is fixed in both cases). One must remove the completing thiol from the mobile phase and use a pure toluene eluant to observe the effect of various organic shell thicknesses on the elution time.

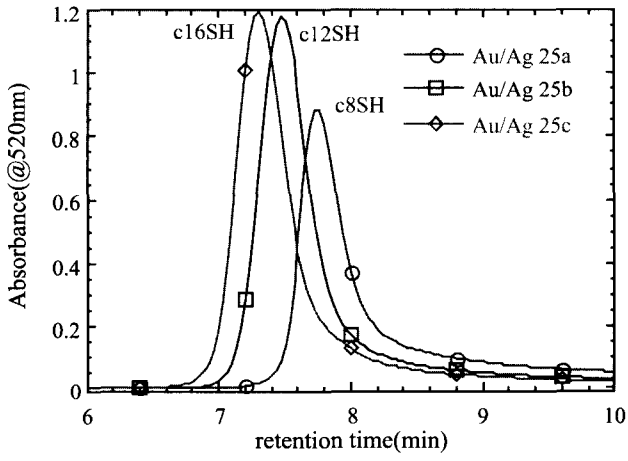


Fig. 15 Effect of organic shell on the total hydrodynamic size of Au/Ag alloy nanoclusters formed by coreduction of the precursor salts in inverse micelles. (Unpublished data.) (*View this art in color at www.dekker.com.*)

What is the effect of alloying or formation of a core–shell structure on the thiol binding behavior? We find that 1:1 (atomic ratio) alloys of a common core size, $D_c \sim 2.0$ nm, elute at distinct times, which correspond to the organic shell thickness as shown in Fig. 15. So the effect of alloying Ag with Au is to prevent ligand exchange. At least two rationalizations of this observation are possible. The change in interatomic coordination and bonding distance in a spherical cluster alloy of Au/Ag affects the thiol metal binding strength, or, more likely, having Au sites at the surface of the metal alloy nanocluster allows the thiol to preferentially bind to these sites. In either case, our alloy AuAg binding results have obvious ramifications for issues in catalysis, where it is empirically known that even small amounts of metal atoms in a majority phase (e.g., Co in MoS_2 hydrodesulfurization catalysts) has a profound effect on the catalytic activity by affecting the Mo substrate binding strength.

Despite the differences in total hydrodynamic size observed in Fig. 15 as a result of the alkyl chain length differences, the optical spectra of each 1:1 Au/Ag alloy nanocluster are indistinguishable, as demonstrated by the elution peak apex absorbance spectra shown in Fig. 16. For similar-sized nanoparticles of pure Ag, the plasmon would occur at 460 nm, while there would be no discernible plasmon peak as a result of damping effects in a pure Au sample of this size. So the effect of alloying on the optical properties of small noble metal nanoclusters is dramatic and constitutes the clearest evidence of a significant composition change from the corresponding pure metal Qdots of equivalent size.

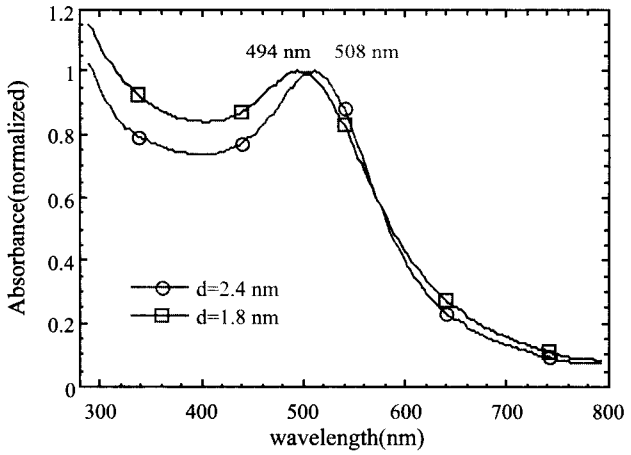


Fig. 17 The peak apex absorbance spectra of a D_c=2.4 and 1.8 nm 1:1 Au/Ag alloy. (Unpublished data.) (*View this art in color at www.dekker.com.*)

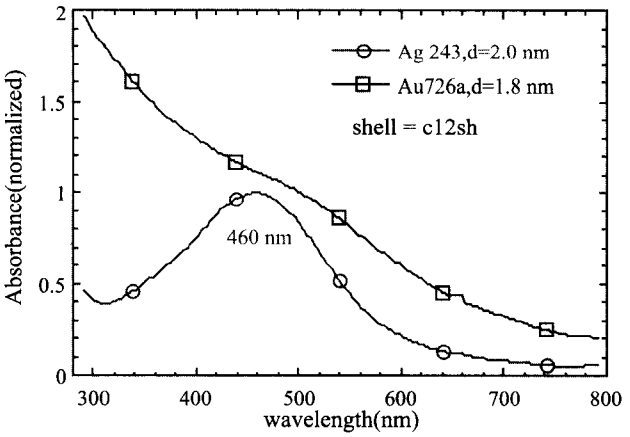


Fig. 18 The peak apex absorbance spectra of $D_c \sim 1.9$ nm pure metal nanoclusters. Compare with the spectra of Fig. 17. (Unpublished data.) (*View this art in color at www.dekker.com.*)

Suppose we grow a 1:1 nanocrystal alloy of a smaller size. Which direction will the plasmon energy shift? If the cluster were pure Ag, a red-shift would be expected, while for pure Au, a blue shift is the norm. The results are shown in Fig. 17. The peak apex absorbance spectra of a 1:1 Au/Ag alloy blue shifts with decreasing size, so the Au component dominates its optical behavior. As expected, an increased damping of the plasmon occurs with decreasing size occurs; however, a distinct absorbance maximum is still observed even for D_c=1.8 nm clusters, apparently because of the silver component, a truly remarkable decrease in the dissipation of the electron plasmon oscillations! A complete theoretical explanation of the alloy behavior in the quantum size regime will prove most challenging, considering that even the extent of damping (peak broadening) and energy shift cannot be predicted currently for pure metal Qdots.

To impress upon the reader how distinct the optical absorbance behavior of the D_c=1.8 nm alloy Qdots are from the pure nanometals of the same size, coated with the identical organic thiol we show in Fig. 18 the spectra from such dots, obtained at the peak apex of the chromatograms.

As one might anticipate, Au/Ag and Ag/Au nanocrystals of the same size have very distinct optical absorbance signatures, which can been easily seen with the naked eye. Consider first the growth of Ag onto a purified seed solution of Au, D_c=6.8 nm, as shown in Fig. 19. The relatively large seed nanocrystals have a distinct absorbance peak at around 520 nm in benzene, which gives them a wine red color. As Ag is slowly deposited onto their surface the absorbance blue shifts, in the direction one would expect for Ag clusters, the peak becomes more symmetrical and narrows with increasing amounts of Ag deposition. In short, the Ag shell component begins to dominate the optical properties.

Fig. 20 shows the case of growth of a Au shell onto a D_c=2.5 nm Ag nanocluster seed. The relatively sharp and symmetrical resonance of the Ag homoatomic cluster broadens and red shifts as Au is deposited. The 1:1 ratio case is worth comparing to the 1:1 alloy of approximately the same size from Fig. 17. One notes that the alloy has both sharper and more red shifted ($\lambda_{max} \sim 508$ nm) compared to the core–shell spectra of Fig. 20. One observes also that the 1:2 Ag/Au nanocluster has not red shifted to the extent one would predict for a pure Au cluster of the same size, but exhibits considerably more dissipation (line broadening) and red-shift than expected for a corresponding homoatomic Ag nanocluster of equivalent size.

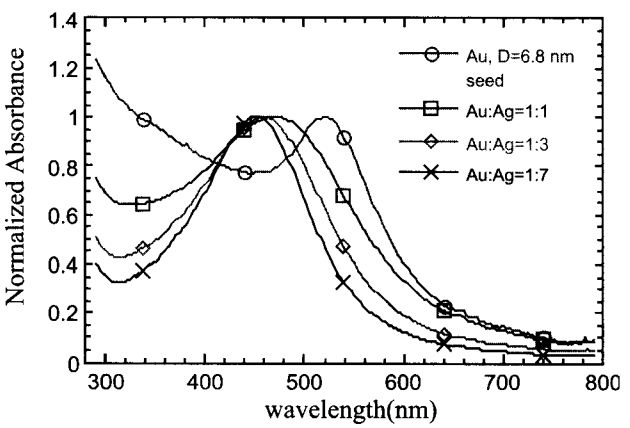


Fig. 19 Optical absorbance for a series of Au/Ag nanoclusters starting from a D_c=6.8 nm Au nanocrystal. The atomic ratios of Au/Ag are indicated in the figure. (Unpublished data.) (*View this art in color at www.dekker.com.*)

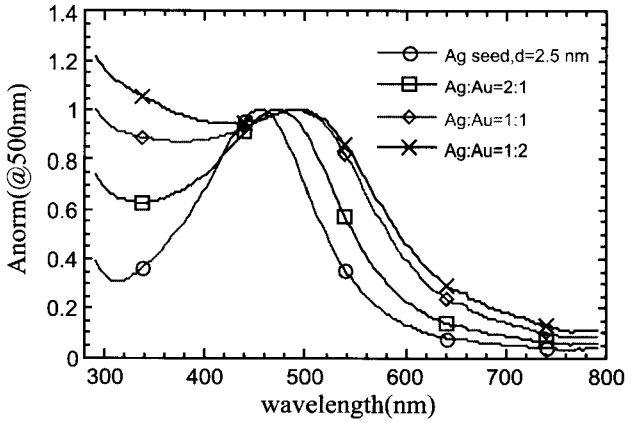


Fig. 20 Optical absorbance for a series of Ag/Au nanoclusters starting from a $D_c = 2.5$ nm Ag nanocrystal. The atomic ratios of Au/Ag are indicated in the figure. (*View this art in color at www.dekker.com.*)

There is a considerable literature discussing thermal decomposition routes to the formation of base metal colloids, which has been recently reviewed.[34] These methods are based on the decomposition of metal carbonyls in the presence of a surfactant or polymer in a relatively high boiling point solvent such as xylene or triethyl benzene. In general, these methods are only effective in producing larger nanoclusters with $D_c > 4$–5 nm with considerable nanocrystalline defects and polydispersity. Recent refinements in these methods have significantly improved on the polydispersity but, unfortunately, because of the rapid nature of the decomposition, significant defects are inevitable and are reflected in the poor magnetic response compared to the bulk counterparts (e.g., ~ 0.2 of M_{sat} for $D_c \sim 6$ nm hcp Co).

What is very interesting is that different structural phases of these nanoclusters may be produced depending on the synthetic approach. We gave the first example of this in a paper describing the effect of surfactant on the formation of either α-Fe (bcc phase) or γ-Fe (high T, fcc phase) using the inverse micelle process.[53] Later work by Sun and Murray[54] showed an even more surprising result for Co, which, when reduced by Superhydride at high T, produced the ε-Co phase, a complex 20-atom/unit cell structure related to β-Mn. These Co nanoparticles could be converted to the hcp phase by annealing at >300°C. However, ε-Co nanocrystalline Co showed significantly lower magnetic response than hcp Co of the same size.

Base Metal Nanoclusters—Co, Fe, and Ni

A puzzle is the observation that, even in the case of directly synthesized nanocrystalline hcp Co, as the size reached the lower size limit of high-T methods, $D_c \sim 3$ nm, a saturation magnetism of only 8 emu/g was observed, about 5% of the bulk value.[34] Even for the largest $D_c \sim 11$ nm Co clusters, a value of only about 60% of the bulk was observed. Likewise, the value of ~60% is about what is the best achieved for nanocrystalline α-Fe. Is the lowered magnetic response attributable to the surface, nanocrystalline defects, spin canting effects, or inadvertent formation of an oxide layer? Although lower density oxide is evident in TEMs of these nanomaterials, its thickness is probably not sufficient to explain the lowered magnetic response, so a likely explanation is poor spin exchange coupling as a result of nanocrystalline defects and/or spin canting near the surface.

By contrast, using the inverse micelle approach, which entails significantly lower nanocrystalline growth rates, we have recently observed saturation magnetism for $D_c \sim 1.8$ nm Co nanoclusters slightly greater than bulk values. This implies full contribution of all the spins to the magnetism, including those at the cluster surface. This is a somewhat unexpected result and seems to depend very sensitively on the chemical nature of the surfactant used to grow the nanoclusters, as well as the time allowed for structural "reconstruction" of the as-synthesized cluster surface. Details will be presented elsewhere.[56] However, our results are consistent with previous observations of very small Co, Fe, and Ni nanoclusters made in cluster beams in vacuum and studied by magnetic deflection combined with mass selection.[24] The fact that magnetic clusters can be grown in solution with high magnetic response and no hysteresis (i.e., they a superparamagnetic at all T) bodes well for potential applications requiring no energy dissipation [transformers and other high radio frequency (RF) inductors].

Formation and Analysis of Quantum Dot Arrays

Hexagonal or cubic packed "rafts" or arrays of quasi-monodisperse spherical colloids [quantum dots arrays (QDAs)] can readily form on various types of substrates upon drying from the liquid phase. Many factors including the initial cluster concentration, the drying rate, the type and physical size of organic ligand "shell," and the cluster core size itself affect the quality of the final cluster film. Just as in bulk crystallization, defects and voids may be introduced if the drying rate is too rapid, trapping colloids in metastable positions. I highly recommend observing the drying of micron-size polystyrene spheres in water on a glass slide under a microscope to understand the role of various factors in the state of the final colloidal film. It is amazing to watch the spheres flow into the drying front and then find the best packing position

(usually hexagonal) for themselves. Sometimes, a sphere will not remain mobile long enough to position itself in the ideal position. Subsequently, this point defect will propagate through the rest of the drying front creating a line defect. In other rare cases, a dumbbell-shaped dimer of two spheres will approach the drying front and, because of its strong intersphere interaction, cannot fit into the proper position in the lattice. Minority spheres of the wrong size and or shape will often be excluded from the lattice, a spontaneous form of size segregation. All of these factors are important on the nanoscale as well although we do not have the ability to observe the dynamics of lattice formation, only the final structures.

Au nanocluster superlattices were first reported in a 1993 paper by Giersig and Mulvaney,[56] who used a modified classical Faraday synthesis in water to produce $\sim$14-nm diameter clusters. Because these samples were charged-stabilized with citrate ions, they were able to employ electrophoretic deposition to form monolayer arrays. Because of the small size of the citrate ion, the gap between particles was fixed and small at $\sim$1 nm.

In 1994, Brustet al.[57] used an inverse micelle process in the presence of alkane thiols to make QDAs of 8-nm Au particles that were cross-linked with a bifunctional dithiol. Robert Whetten's group at Georgia Tech was the first to demonstrate the formation of QDAs with relatively small Au nanoparticles produced by using the inverse micelle technique with large amounts of thiol present during the reduction to severely limit the cluster growth. They also showed 3-D superlattices of significant ($\sim$1–10 μm) size.[58] Later work by Heath and coworkers utilizing inverse micelle techniques produced highly polydisperse dedecanthiol capped Au clusters which still formed ordered arrays in which they noted that size exclusion of smaller clusters occurred spontaneous during the drying process.[59]

In 2000, we reported the first quantitative studies of the effect of particle size and the organic capping thickness on the spacing and degree of ordering in Au QDAs.[41] These studies were enabled by the development of the first detailed digital image analysis methods for metallic nanocluster QDAs. Later, we reported similar studies on Pt QDAs, which had not been made previously.[42] We will summarize highlights of those papers now.

First, the formation and, in particular, the domain size of QDAs depend on many experimental factors. The successful formation of superlattices depends on the proper choice of oil, surface passivant, metal ligand binding strength, ligand chain length or structure, substrate type, and rate of drying.

Consider, for example, the effect of metal ligand binding strength on QDA stability. We have demonstrated via SEC that alkyl thiols bind considerably more strongly to Au and Pt nanoclusters of a fixed size than they do to Ag. This is reflected in the stability under TEM vacuum conditions of Au arrays compared to Ag ones. Although both initially form large, highly ordered hexagonal domains as observed in TEM, Ag QDAs rapidly (<1 day) deteriorate as a result of alkyl thiol desorption followed by cluster fusion or sintering.

We have also shown using SEC that the alkyl chain length affects the binding affinity of metallic Qdots for thiol surfactants. We find that longer chain length thiols bind more strongly, for example, to Au nanoclusters. TEM studies of QDAs, as a function of alkyl chain length k, confirm that short chain, $k<8$ thiols, fail to form large ordered domains. What is more surprising is that an optimal value, $k=10$–14, is observed, with worse array formation despite stronger binding affinity for $k=16$. Perhaps this is attributable to the observation that alkane ligands interdigitate and facilitate the QDA formation, and $k=16$ is approaching the length at which this process is still efficient.

The rate of drying and concentration of the initial cluster solution is also important in determining the structure of the final QDA. The effects here are a little more obvious, namely, faster drying and lower initial concentration favors the formation of monolayers of

Fig. 21 Optical micrograph of QDA 3-D superlattices made of Au nanocrystals with $D_c \sim 4.6$ nm. (From Ref. [41].)

hexagonally or square packed Qdots, while higher concentrations and slower drying (less-volatile solvents) favors cluster pile-up and formation of bilayers and 3-D crystals on the substrate. In fact, 3-D crystals of Ag form so readily upon slow drying from solution that they can be observed on the glass surfaces of vials with inadequately sealed lids. Naturally, such molecular crystals dissolve readily into organic solvents.

Finally, we found that use of substrates such as Teflon was more effective than glass for the formation of large ($\sim$10–100 μm), supracrystals. An example (Fig. 21) is shown below. The triangular facets or crystal habit adopted by these QDA superlattices seems to depend on the nanocrystal core size, but this is a very complex issue (see below). Graphite proved to be too porous for good QDA crystal superlattice formation, and holey carbon TEM grids, which are also relatively porous were more effective for the formation of either mono- or bilayer QDAs. However, these general observations are also particular to the metal type, with Ag almost always showing both mono- and multilayer formation using holey carbon grids. One important technical point is that, for hydrophilic substrates such as glass, inert grease (Krytox) should be used to contain the nonwetting oil.

The nanocrystal size-dependent habit adopted by 3-D QDAs is amusingly illustrated in Fig. 22, for D_c=4.0 nm Pt nanocrystals.[42] The "Mitsubishi" motif exhibited by these crystals is in marked contrast to that exhibited by smaller, D_c=2.5 nm crystals in Fig. 23. In both cases, dodecanthiol was used to passivate the surface, decane was used as the solvent, and an identical Pt concentration and substrate was used.

On larger-length scales still the QDAs of Fig. 21 form dentritic films, Fig. 24, of "cluster matter," whose complex optical and electrical behavior was the subject of a fairly recent paper and is discussed in Kreibig and Vollmer's monograph.[2] These dentritic films were shown to sinter into a metallic films with high optical reflectivity and electrical conductivity at low temperatures ($\sim$200°C).[60]

Fig. 23 An optical micrograph is Pt QDAs formed by very small (D_c=2.5 nm) nanocrystals with slightly irregular shapes shows that crystal habit depends on the nanocrystal core size. (From Ref. [42].)

Finally, it should be mentioned that, depending on growth conditions, many other crystal habits were observed, including large ($\sim$1–10 μm) hexagonal motifs. The conditions that favor one habit over another are just as complex and difficult to understand as that of macrocrystalline growth from supersaturated solutions, and growing QDAs of large size and high perfection is very much an art. An obvious goal would be to obtain crystals of sufficient size to perform single-crystal XRD and determine each atomic position in each identical nanocrystal to atomic resolution. This may well be possible in

Fig. 22 A tri-foil crystal motif is observed in an optical micrograph of D_c=4.0 Pt nanocrystals. It is reminiscent of a well-known Japanese corporate logo (although no funding was provided to influence this fortuitous observation). (From Ref. [42].)

Fig. 24 Optical micrograph at low magnification of the triangular Au QDAs shown in Fig. 21. (From Ref. [41].)

the future. It would answer some currently unknowable issues with respect to surface structure in nanocrystals with major impact on the design of nanometal catalysts, nanophosphors, etc.

TEM image analysis

The qualitative observations of the preceding section can be augmented in the case of QDA monolayers by employing automated digital image analysis of large areas of TEM grids. Consider the TEM of Fig. 25. We first determine the center-of-mass position, and average size of all acceptable (i.e., nonoverlapping) nanocrystals in the image. These are then replaced with digital, color-coded balls of the same size, and the particle coordinates can be subsequently used to calculate both the 2-D and radial-averaged, $g(r)$, correlation functions. The latter is the Fourier transform of $S(q)$, the scattering structure factor which would be obtained from, for example, small-angle X-ray scattering (SAXS)[44] or neutron scattering (SANS). Balls that are sufficiently out of register with neighboring balls (e.g., in a hexagonal lattice) are then assigned colors that identify the coherent domains (analogous to the peak linewidth in SAXS). Fig. 26 shows the result of such assignments.

With the available domain and nanocrystal position information, one can then rotate each domain by the correct amount to obtain a coherent domain for the entire region (balls which fail to belong to any of the domains are eliminated for simplicity). The resulting rotated 2-D correlation function is shown in Fig. 27. One easily observes the hexagonal symmetry of the QDA in this figure, and the very high degree of long-range order.

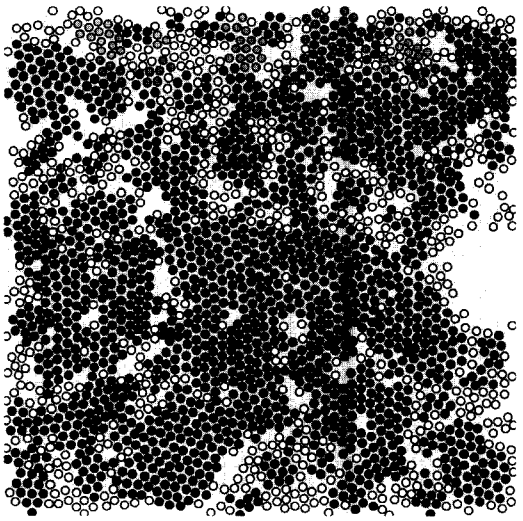


Fig. 26 The digitized ball file containing the center-of-mass coordinates of the nanocrystals and whose ball color codes for the domain of a given ball is shown. (From Ref. [41].) (*View this art in color at www.dekker.com.*)

A radial average can then be performed to obtain $g(r)$ and the average interparticle spacing. This is shown in Fig. 28. The hexagonal symmetry of the lattice is reflected in the splitting of the second peak, which is harder to observe when the domains are incoherent (i.e., not rotated). The average interparticle spacing and the gap can be very precisely determined from such analysis. This allowed us to perform a very precise determination of the effect of alkyl chain length (organic shell thickness) on the

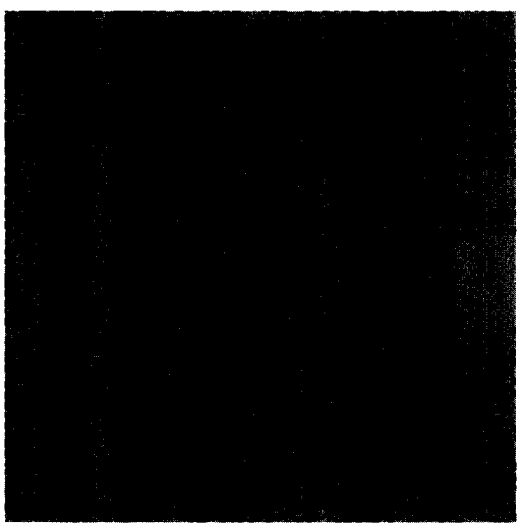


Fig. 25 A monolayer of D_c=4.6 nm Au clusters that forms 3-D crystals as shown in Fig. 21. (From Ref. [41].)

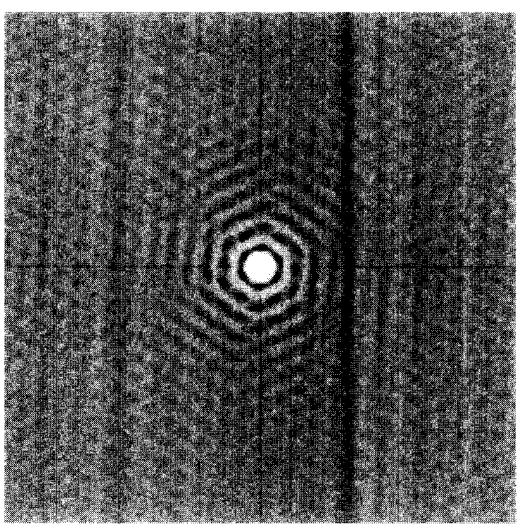


Fig. 27 Rotated 2-D pair correlation function from Fig. 26. (From Ref. [41].)

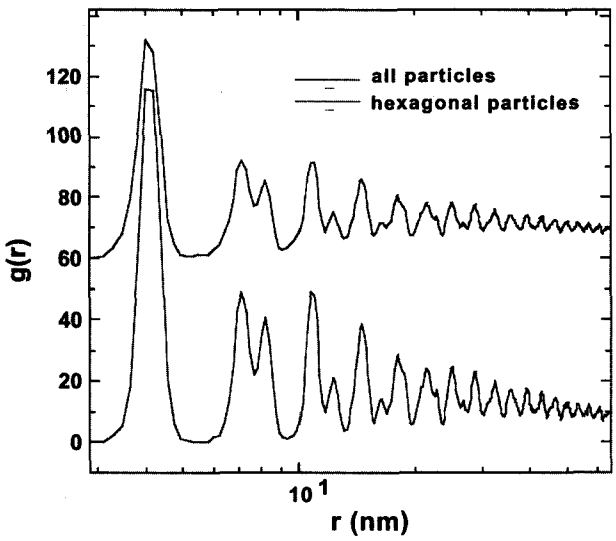


Fig. 28 The radial pair correlation function, $g(r)$ reflects both the degree of order (the number of periods or oscillations) and the symmetry of the QDA lattice. (From Ref. [41].) (*View this art in color at www.dekker.com.*)

interparticle spacing in Au and Pt QDAs for the first time. We reported a change of ~1.2 Å/ carbon for alkyl chains Ck with $6<k<14$, in the case of Au and 1.4 Å/carbon in the case of Pt QDAs. For comparison, SEC determination of the organic shell thickness in Au alkylthiolate nanoconjugates found an increase of 2.4 Å/2 carbon atoms in the shell thickness in toluene—a very good agreement. It is important to note that substantial interdigitation of the alkyl groups is always observed in the case of Au, Ag, Pt, and Pd QDAs, and plays an important role in the formation of the arrays. In the case of surfactants that do not easily pack or are too short, QDA formation is very rare. Also, for particles smaller than ~2.5 nm, long-range order is rarely observed and the particles prefer to disperse randomly on the grid.

The symmetry of a monolayer lattice obtained from a given nanocluster solution is typically hexagonal; however, in the same preparation on the same grid, we have also observed and reported square lattice packings, so the symmetry cannot be solely dependent on organic layer thickness, nanocluster metal type, nor solvent because these are all identical. It does indicate that these two lattice types are of very similar energy.

Similarly, unexpected packing can occur in bilayers, as demonstrated by Figs. 29 and 30. In Fig. 29, a Pt PDA bilayer we observe the "expected" low-energy trigonal (threefold) locations of the second layer above the "holes" in the first hexagonal layer. But in other Pt QDA, Fig. 30, we observed that clusters in the second layer occupy the twofold sites.

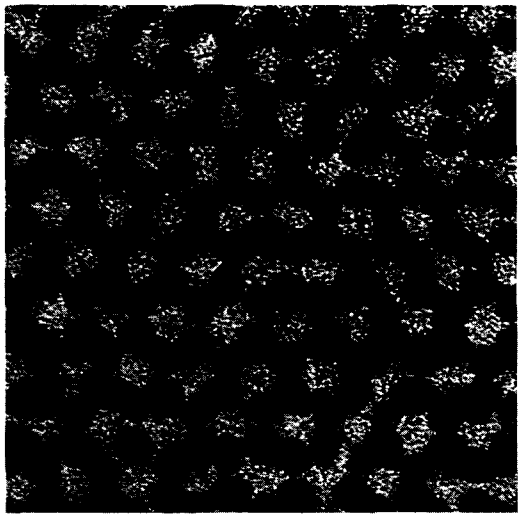


Fig. 29 Pt QDA bilayer with the clusters deposited onto the low-energy trigonal sites above the first monolayer. (Unpublished data.)

Nanocluster Matter and Nanotechnology

Cluster matter is obviously significantly more complex to study and understand than dilute solutions of individual Qdots acting independently. However, it does present some intriguing technological possibilities. Consider, for example, the effect of nanocrystal diameter and particle spacing on the effective dielectric constant of a fcc (or hexagonal) QDA. Fig. 31 shows that as the ratio of D_c/gap increases, the dielectric constant can become very large. This has implications for energy storage in capacitors

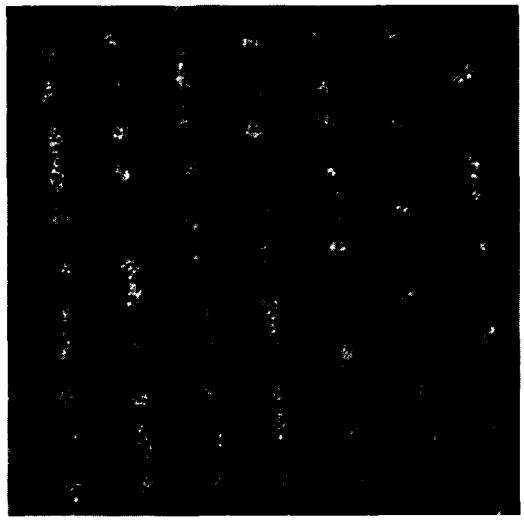


Fig. 30 Pt Qdot bilayer has the second layer of clusters deposited onto twofold symmetry sites. (Unpublished data.)

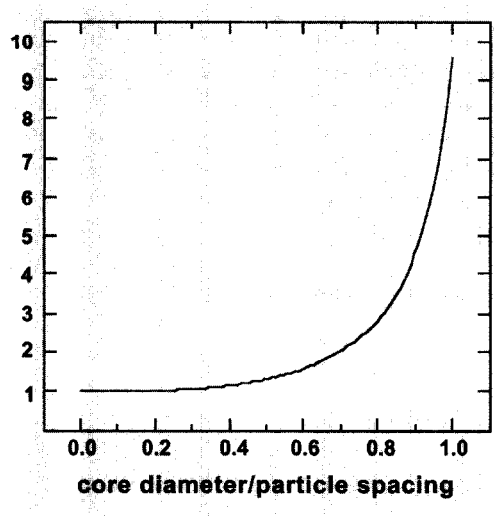


Fig. 31 The calculated dielectric constant of an fcc QDA superlattice increases dramatically with core diameter/gap ratio. (From Ref. [41].) (*View this art in color at www.dekker.com.*)

because cluster matter may not suffer the same problems with increased dielectric breakdown as those known for bulk powders. We discuss the reasons for these difference in Ref. [41].

Another area in which metallic nanoclusters are likely to have a significant impact in the future is in heterogeneous catalysis. As examples, we cite our work on nanosize Fe and FeS_2 as coal hydrogenation and hydrogenoloysis catalysts, where the nanoclusters are directly deposited onto the coal powder, which acts both as a support and a substrate.[61,62] The coal is broken down under high-pressure H_2 into liquids, which may serve as either chemical feedstocks for advanced engineering materials or as fuels in energy production. A very significant amount of the chemical products are aromatic or napthalenic in structure, a very desirable starting point for many engineering polymers and plastics.

We have also investigated nanosize MoS_2 as an photocatalyst demonstrating a size-dependent photooxidation of both phenol[50] and pentachlorophenol.[51] The latter is a major environmental contaminant as it has been widely employed as a fungacide and herbacide for wood preservation. As MoS_2 has a two-dimensional graphite-like structure with the catalytically active Mo sites at the "edge" of the cluster, making this material nanosize vastly increases its potential catalytic activity. Already, it is the most useful fuel hydrodesulfurization catalyst known. (This is the critical first step in crude refining, in which heteroatom compounds are removed to avoid poisoning subsequent expensive transition metal catalysts.) One would anticipate that nanosize MoS_2 would be even more effective (perhaps at lower temperature and H_2 pressures).

We have examined nanosize Pd, Pt, and Rh deposited on high surface area, commercially available supports such as alumina and carbon and found that not only can milder T and P conditions be utilized for hydrogenation, but that the product distribution (selectivity) can be tuned by nanocluster size and metal type. In fact, we have discovered that Pd nanoclusters on carbon can actually be more effective than more expensive Rh in the hydrogenation of pyrene. This is not true of commercially available Pd and Rh catalysts, and indicates the almost limitless possibilities in this largely unexplored area of nanotechnology.

CONCLUSION

In this review, we discussed chemical preparation methods for the synthesis of metallic quantum dots, emphasizing the important role of surface active agents for the control of the average cluster size and size dispersion. Inverse micelle synthesis in nonpolar oils using readily available, inexpensive ionic metal salts encapsulated in their interior is the most versatile method for metallic Qdot synthesis. We demonstrated the use of liquid chromatography for cluster size analysis and the role such rapid feedback plays in synthesis development. We demonstrated that special, thermodynamically stable cluster sizes are formed in the size regime $D_c<3$ nm, that the binding strengths of ligands to metal nanoclusters is stronger for Au than Ag, and that the length of the thiolate itself affects the binding affinity. We noted that short chain length alkyl thiolates may act at etchants to narrow the size distribution with time, and that, in general, an evolution of the cluster size distribution does occur, its rate being correlated with the bulk melting T of the material (i.e., slower evolution for Pt than for Au or Ag).

We next reviewed the size-dependent optical properties of Au and Ag nanoclusters, optical properties of core–shell particles, and the unexpected visible and NIR photoluminesence from molecular sized nanometals. A blue shift of the plasmon resonance was observed for nanosize Au, while the opposite behavior was found for Ag and attributed the greater role of inter-d-band transitions in Au than in the case of Ag. We reported the emergence of distinct molecule-like absorbance features in Au clusters approaching 1 nm in dimension. The optical properties of core–shell particles of equivalent size and atomic composition were shown to depend on whether Ag or Au was in the interior and were further shown to depend on whether an alloy or a true core–shell type structure was synthesized.

We next discussed the various synthetic approaches and resulting magnetic properties of the base itinerant metals, Co, Fe, and Ni. We discussed the possible reasons for the heretofore low magnetic response of nanocrystals of these materials and also discussed soon-to-be published work on ultrasmall, $D_c = 1.4$ and 1.9 nm Co nanocrystals exhibiting greater-than-bulk magnetic response. Possible reasons connected with surface magnetism and surface reconstruction in nanocrystals were given as explanations of our observations. We also discussed the synthesis of nonequilibrium crystal structures of nanoparticles of Fe and Co.

We next discussed the formation of QDAs of nanocrystals of Au, Pt, and Pd outlining the important parameters controlling the degree of order and whether mono-, bi-, or multilayer QDAs form. In particular, we reviewed novel digital-image analysis methods for obtaining the average domain size, interparticle spacing, and extent of long-range order in TEM images of QDAs. We noted that bilayers may exhibit either the predicted low-energy, threefold coordination or a twofold coordination in the second layer. We also commented that either hexagonal or square packing of nanocrystals can occur in the same sample on the same substrate.

We concluded our review with a discussion of the relationship between Qdots and nanotechnology. The ability to control the size, shape, and interface structure of Qdots is critical to their application in nanotechnology as new magnetic, dielectric, optical, and catalytic materials. As one example, we discussed the effect of core size to interparticle ratio on the dielectric constant and pointed out that one may expect nanosize clusters to eliminate some of the dielectric breakdown issues found in macroscopic materials, allowing significant increases in energy storage in capacitors. We particularly emphasized catalytic applications because we feel these may yield the most spectacular improvements on conventional technology. Two examples, the direct conversion of coal to chemicals and fuels using Fe and FeS_2, and solar-driven photocatalysis using MoS_2 were given.

In any review, one must be quite selective in the topics covered. But we hope to have given the reader a good overview of the reasons for the large scientific interest in nanomaterials and nanotechnology, as well as the tools which are currently being developed to further these fields. Essentially, scientists are redoing all of materials research on a length scale in which interface is dominant. Thus these endeavors are likely to continue for many years and yield many unexpected observations.

ACKNOWLEDGMENTS

This work was supported by the Division of Materials Sciences, Office of Basic Energy Sciences of the US Department of Energy under contract DE-AC04-94AL8500. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the Department of Energy. The author would like to acknowledge the support of many collaborators over the years, especially my technical assistants, Edward Vernon, Sharon Craft, Judy Odinek, and James Mikkaelson. I would also like to recognize the fine TEM, XRD, and SAD work of Paula Provencio and my long-time friend and collaborator James Martin.

REFERENCES

1. Faraday, M. Experimental relations of gold (and other metals) to light. Philos. Trans. R. Soc. Lond. **1857**, *147*, 145–185.
2. Kreibig, U.; Vollmer, M. *Optical Properties of Metal Clusters*; Springer-Verlag: Berlin, 1995.
3. Hoon, S.R.; Kilner, M.; Russell, G.J.; Tanner, B.K. Preparation and properties of nickel ferrofluids. J. Appl. Polym. Sci. **1966**, *10*, 1915–1927.
4. Hess, P.; Parker, P.S., Jr. Polymers for stabilization of colloidal cobalt particles. J. Appl. Polym. Sci. **1966**, *10*, 1915–1927.
5. Griffiths, C.H.; O'Horo, M.P.; Smith, T.W. The structure, magnetic characterization, and oxidation of colloidal iron dispersions. J. Appl. Phys. **1979**, *50*, 7108–7115.
6. Murray, C.B.; Norris, D.J.; Bawendi, M.G. Synthesis and characterization of nearly monodisperse CdE (E=S, Se, Te) semiconductor nanocrystallites. J. Am. Chem. Soc. **1993**, *115*, 8706–8715.
7. Schmid, G. Hexachlorodecakis(triphenyl-phosphine)pentrapentacontagold, $Au_{55}[P(C_6H_5)_3]_{12}C_{16}$. Angew. Chem. **1978**, *90*, 417–420.
8. Schmid, G. *Developments in Transition Metal Cluster Chemistry—The Way to Large Clusters. Structure and Bonding*; Springer-Verlag: Berlin, 1985; Vol. 62, 52–82.
9. Boutonnet, M.; Kizling, J.; Stenius, P. The Preparation of monodisperse colloidal metal particles from microemulsions. Colloids Surf. **1982**, *5*, 209–225.
10. Boutonnet, M.; Kizling, J.; Mintsa-Eya, V.; Choplin, A.; Touroude, R.; Maire, G.; Stenius, P. Monodisperse colloidal metal particles from nonaqueous solutions: Catalytic behavior in hydrogenation of but-1-end of platinum, palladium, and rhodium particles supported on pumice. J. Catal. **1987**, *103*, 95–104.
11. Wilcoxon, J.P.; Williamson, R.L. *Formation of Metal Colloids in Inverse Micelles and Microemulsions*; Proceedings of the MRS, 1989, Boston, MA.

12. Wilcoxon, J.P. Method for the Preparation of Metal Colloids in Inverse Micelles and Product Preferred by the Method. U.S. Patent #5,147,841, Sep. 15, 1992, issued.
13. Wilcoxon, J.P.; Williamson, R.L.; Baughman, R.J. Optical properties of gold colloids formed in inverse micelles. J. Chem. Phys. **1993**, *98* (12), 9933–9943.
14. Wilcoxon, J.; Sylwester, A.; Nigrey, P.; Martino, A.; Quintana, C.; Baughman, R. *Formation and Characterization of Highly-Dispersed Iron Catalysts*; Eighth Annual International Pittsburgh Coal Conference Proceedings, 1991; 703–708.
15. Wilcoxon, J.P.; Martino, R.I.; Baughmann, R.L.; Klavetter, E.; Sylwester, A.P.*Synthesis of Transition Metal Clusters and their Catalytic and Optical Properties*; Proceedings of the Fall MRS, Symposium J, Boston, MA, 1992.
16. Wilcoxon, J.P.; Martino, T.; Klavetter, E.; Sylwester, A.P. Synthesis and Catalytic Properties of Metal and Semiconductor Nanoclusters. In *Nanophase Materials*; Hadjipanaysis, Siegel, Eds.; 1993; 770–780.
17. Martino, A.; Wilcoxon, J.P.; Kawola, J. Synthesis and characterization of Fe and FeS_2 (pyrite) coal liquefaction catalysts in inverse micelles. Energy Fuels **1994**, *8*, 1289–1294.
18. Wilcoxon, J.P.; Samara, G.A.; Provencio, P.N. Optical and electronic properties of Si nanoclusters synthesized in inverse micelles. Phys. Rev., B **1999**, *60*, 2704–2714.
19. Wilcoxon, J.P.; Samara, G.A. Tailorable, visible light emission from silicon quantum dots. Appl. Phys. Lett. **1999**, *21*, 3164–3166.
20. Wilcoxon, J.P.; Samara, G.A.; Provencio, P. Synthesis and optical properties of colloidal germanium nanocrystals. Phys. Rev., E **2001**, *64*, 35417-1.
21. de Heer, W.A. Simple metal clusters. Rev. Mod. Phys. **1993**, *65*, 612–645.
22. Brack, M. The physics of simple metal clusters: Self-consistent jellium model and semiclassical approaches. Rev. Mod. Phys. **1993**, *65*, 677–702.
23. Friedlander, G.; Kennedy, J.W.; Miller, J.M. *Nuclear and Radiochemistry*, 2nd Ed.; John Wiley & Sons: New York, 1966.
24. Billas, I.M.L.; Chatelain, A.; De Heer, W.A. Magnetism of Fe, Co and Ni clusters in molecular beams. J. Magn., Magn. Mater. **1997**, *168* (1–2), 64–84.
25. Mie, G. Ann. Phys. **1908**, *25*, 377.
26. Wilcoxon, J.P.; Samara, G.A. Strong quantum-size effects in a layered semiconductor: MoS_2 nanoclusters. Phys. Rev., B Rapid Commun. **1995**, *51*, 7299–7302.
27. Wilcoxon, J.P.; Samara, G.A.; Newcomer, P. Strong quantum confinement effects in semiconductors: FeS_2 nanoclusters. Solid State Commun. **1996**, *98*, 581–585.
28. Wilcoxon, J.P.; Newcomer, P.; Samara, G.A. Synthesis and optical properties of MoS_2 and isomorphous nanoclusters in the quantum confinement regime. J. Appl. Phys. **1997**, *81*, 7934–7940.
29. Wilcoxon, J.P.; Provencio, P. Use of surfactant micelles to control the structural phase of nanosize iron clusters. J. Phys. Chem., B **1999**, *103*, 9809–9813.
30. Wilcoxon, J.P.; Martin, J.E.; Provencio, P. Size distributions of gold nanoclusters studied by liquid chromatography. Langmuir **2000**, *16* (25), 9912–9920.
31. Wilcoxon, J.P.; Provencio, P. Etching and aging effects in nanosize Au clusters investigated using high resolution size-exclusion chromatography. J. Am. Chem. Soc. **2002**, submitted.
32. Brust, M.; Fink, J.; Bethell, D.; Schriffin, D.J.; Kiely, C. Synthesis and reactions of functionalised gold nanoparticles. J. Chem. Soc., Chem. Commun. **1995**, 1655–1658.
33. Whetten, R.L.; Khoury, J.T.; Alvarez, M.M.; Murthy, S.; Vezmar, I.; Wang, Z.L.; Stephens, P.W.; Clevelend, C.L.; Luedtke, W.D.; Landman, U. Nanocrystal gold molecules. Adv. Mater. **1996**, *8*, 428–437.
34. Murray, C.B.; Sun, S.; Coyle, H.; Betley, T. Monodisperse 3d transition-metal (Co, Ni, Fe) nanoparticles. MRS Bull. **2001**, *26* (12), 985–990.
35. Wilcoxon, J.P.; Craft, S.A. Liquid chromatographic analysis and characterization of inorganic nanoclusters. NanoStructur. Mater. **1997**, *9*, 85–88.
36. Wilcoxon, J.P.; Martin, J.E.; Provencio, P. Optical properties of gold and silver nanoclusters investigated by liquid chromatography. J. Chem. Phys. **2001**, *115*, 998–1008.
37. Yau, W.W.; Kirkland, J.J.; Bly, D.D. *Modern Size-Exclusion Liquid Chromatography*; Wiley-Interscience: New York, 1979.
38. Norris, D.J.; Sacra, A.; Murray, C.B.; Bawendi, M.G. CdSe-TOP coated nanocrystals display band gap emission with room—T Q.Y of a few percent. Phys. Rev. Lett. **1994**, *101*, 8455–8459.
39. Dabbousi, B.O.; Rodriguez-Viejo, J.; Mikulec, F.V.; Heine, J.R.; Mattoussi, H.; Ober, R.; Jensen, K.F.; Bawendi, M.G. (CdSe)ZnS core–shell quantum dots: Synthesis and characterization of a size series of highly luminescent nanocrystallites. J. Phys. Chem., B **1997**, *101*, 9463–9474.
40. Peng, X.; Schlamp, M.C.; Kadavanich, A.V.; Alivisatos, A.P. Epitaxial growth of highly luminescent CdSe/CdSe core/shell nanocrystals with

Q

photostability and electronic accessibility. J. Am. Chem. Soc. **1997**, *119*, 7019–7029.
41. Martin, J.E.; Wilcoxon, J.P.; Odinek, J.; Provencio, P. Control of the interparticle spacing in gold nanoparticle superlattices. J. Phys. Chem., B **2000**, *40*, 9475–9486.
42. Martin, J.E.; Wilcoxon, J.P.; Odinek, J.; Provencio, P. Superlattices of platinum and palladium nanoparticles. J. Phys. Chem., B **2002**, *106*, 971–978.
43. Schaaff, T.G.; Shafigullin, M.N.; Khoury, J.T.; Vezmar, I.; Whetten, R.L.; Cullen, W.G.; First, P.N.; Gutierrex-Wing, C.; Ascnsio, J.; Jose-Yacaman, M.J. Isolation of smaller nanocrystal Au molecules: Robust quantum effects in optical spectra. J. Phys. Chem., B **1997**, *101*, 7885–7892.
44. Wilcoxon, J.P.; Thurston, T.R.; Craft, S. Performance of a high resolution Bonse–Hart SAXS instrument using a synchrotron source. Rev. Sci. Instrum. **1996**, *67* (9), 3021–3034.
45. Schaaff, T.G.; Whetten, R.L. Controlled etching of Au:SR cluster compounds. J. Phys. Chem., B **1999**, *103*, 9394–9396.
46. Link, S.; Beeby, A.; FitzGerald, S.; El-Sayed, M.A.; Schaaff, T.G.; Whetten, R.L. Visible to infrared luminescence form a 28-atom gold cluster. J. Phys. Chem., B **2002**, *106*, 3410–3413.
47. Wilcoxon, J.P.; Provencio, P. In Optical Properties of II–VI Semiconductor Nanoclusters for Use as Phosphors. Proceedings of the SPIE conference on "Opitcal Properties of Nanocrystals", Seattle, WA, July 7–11, 2002.
48. Wilcoxon, J.P.; Martin, J.E.; Parsapour, F.; Wiedenman, B.; Kelley, D.F. Observation of photoluminescence from nanosize gold clusters. J. Chem. Phys. **1998**, *108*, 9137–9143.
49. Kreibig, U.; Genzel, L. Optical absorption of small metallic particles. Surf. Sci. **1985**, *156*, 678–682.
50. Thurston, T.R.; Wilcoxon, J.P. Photo-oxidation of organic chemicals catalyzed by nanoscale MoS_2. J. Phys. Chem. **1998**, *103*, 11–17.
51. Wilcoxon, J.P. Catalytic photoxidation of pentachlorophenol using semiconductor nanoclusters. J. Phys. Chem. **2000**, *104*, 7334–7343.
52. Wilcoxon, J.P.; Provencio, P. Heterogeneous nucleation and growth on metal nanoclusters. Science, in preparation.
53. Wilcoxon, J.P.; Newcomer, P. Use of surfactant micelles to control the structural phase of nanosize iron clusters. J. Phys. Chem., B **1999**, *103*, 9809–9812.
54. Sun, S.; Murray, C.B. Synthesis of monodisperse cobalt nanocrystals and their assembly into magnetic superlattices (invited). J. Appl. Phys. **1999**, *85*, 4325–4329.
55. Wilcoxon, J.P.; Venturini, E.L.; Provencio, P. Enhanced magnetic response of dilute cobalt nanoparticles in an organic matrix—The effects of aging and interface structure. Phys. Rev. Lett., in preparation.
56. Goersig, M.; Mulvaney, P. Formation of ordered two-dimensional gold colloid lattices by electrophoretic deposition. J. Phys. Chem. **1993**, *97*, 6334–6336.
57. Brust, A.; Berthell, D.; Schriffrin. D.J.; Kiely, C.J. Novel gold-dithiol nano-networks with non-metallic electronic properties. Adv. Mater. **1995**, *7*, 795–797.
58. Fink, J.; Kiely, C.J.; Bethell, D.; Schiffrin, D.J. Self-organization of nanosized gold particles. Chem. Mater. **1998**, *10*, 922–925.
59. Taleb, A.; Petit, C.; Pileni, M.P. Optical properties of self-assembled 2D and 3D superlattices of silver nanoparticles. J. Phys. Chem., B **1998**, *102*, 2214–2220.
60. Martin, J.E.; Odinek, J.; Wilcoxon, J.P.; Anderson, R.A.; Provencio, P. Sintering of gold and platinum nanoclusters. J. Phys. Chem. B **2002**, *in press.*
61. Wilcoxon, J.; Sylwester, A.; Nigrey, P.; Martino, A.; Quintana, C.; Baughman, R. *Formation and Characterization of Highly-Dispersed Iron Catalysts*; Eighth Annual International Pittsburgh Coal Conference Proceedings, 1991; 703–708.
62. Martino, A.; Wilcoxon, J.; Sylwester, A.P.; Kawola, J.S. *Synthesis and Characterization of Fe and FeS_2 (pyrite) Catalysts Particles in Inverse Micelles*; Proceedings of the 205th National American Chemical Soc. Meeting, 1993.

Quantum Dots: Phonons in Self-Assembled Multiple Germanium Structures

Q

Jianlin Liu
University of California, Riverside, California, U.S.A.

Aleksandr Khitun
Kang L. Wang
University of California, Los Angeles, California, U.S.A.

INTRODUCTION

Self-assembled Ge quantum dots by the Stranski–Krastanov growth mode have attracted much attention for many years. Similar to the purpose of the research on short-period Si/Ge superlattices and Er-doped Si, self-assembled Ge/Si quantum dots may be exploited to fabricate Si-based on-chip light emitting sources for 1.55-μm fiberoptic communication applications. Takagahara and Takeda[1] and Ren[2] have theoretically predicted that an indirect-to-direct conversion of the optical transition of SiGe quantum dots would occur whenever the sizes of the quantum dots were small enough. In order to increase the component of quasi-direct transition of the Ge/Si quantum dot system and thus to enhance light emission intensity of the quantum dots for practical applications, the understanding of nonradiative recombination mechanisms (such as phonon-assisted process) is essential.

To probe phonons from self-assembled semiconductor quantum dots, Raman spectroscopy is an efficient and indispensable experimental tool. Up to date, several groups have already reported Raman scattering studies of self-assembled quantum dots of group III–V systems, such as (In, Ga, Al)Sb/GaAs,[3] InSb/InP,[4] In(Ga)As/GaAs,[5–9] InAs/InP,[10–13] (Al, Ga)As/InAs,[14] InAs/AlAs,[15,16] and GaN/(Si)AlGaN,[17,18] group II–VI system, such as CdSe/ZnSe,[19] and group IV system, i.e., Ge/Si.[20–38] Optical phonon spectra of any of these systems were basically used to extract the chemical composition of the quantum dots as a result of interdiffusion between the dots and the surrounding media or the substrates. It should be pointed out that a few other techniques for the determination of composition of self-assembled quantum dots have been reported. These include scanning tunneling microscopy,[39] transmission electron microscopy (TEM) with high-resolution imaging,[40] electron energy loss spectrometry,[41] X-ray energy disperse spectrometry,[42] high-resolution X-ray diffraction,[43,44] and scanning TEM.[45] Most of these techniques are capable to show nonuniform dot material distribution in the dots. Optical phonon Raman scattering method is simple and direct to give an average concentration of the dots. In contrast to the tremendous research on optical phonons, the effort on the research of phonon process in the low-frequency acoustic spectral region, however, is much smaller. In self-assembled Ge quantum dot system, for example, one early work reported the observation of equal-distance acoustic peaks in 25-period Ge quantum dot superlattices.[21] Afterwards, Milekhin et al. investigated folded longitudinal acoustic phonons in their Ge dot superlattices and explained the acoustic vibrations by the elastic continuum model.[26–29] Recently, resonant Raman scattering by acoustic phonons in double- and multilayered Ge dot structures was reported and the observed equal-distance oscillation peaks were explained by interference and ordering effects.[32,33] The origins of these observed low-frequency acoustic phonon spectra in Ge dot superlattices therefore remain unclear and debatable because there is a lack of systematic studies, such as the dependence of acoustic phonons on island sizes and other island-related parameters.[23]

In this entry, we systematically study Raman scattering by optical and acoustic phonons in multiple Ge quantum dots. The analysis of GeGe and SiGe optical phonon features takes the phonon confinement effect, strain effect, and atomic intermixing into account. Acoustic phonons are found to originate from folded acoustic phonons associated with a superlattice and can be explained by elastic continuum model.

RESULTS ON SELF-ASSEMBLED MULTIPLE GE QUANTUM DOTS

All samples were grown on Si (100) wafers using a solid source molecular beam epitaxy (MBE) system. Samples A, B, and C were grown at 540°C with 10 periods of Ge and Si bilayers and contained a Ge coverage of 12, 15, and

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013529

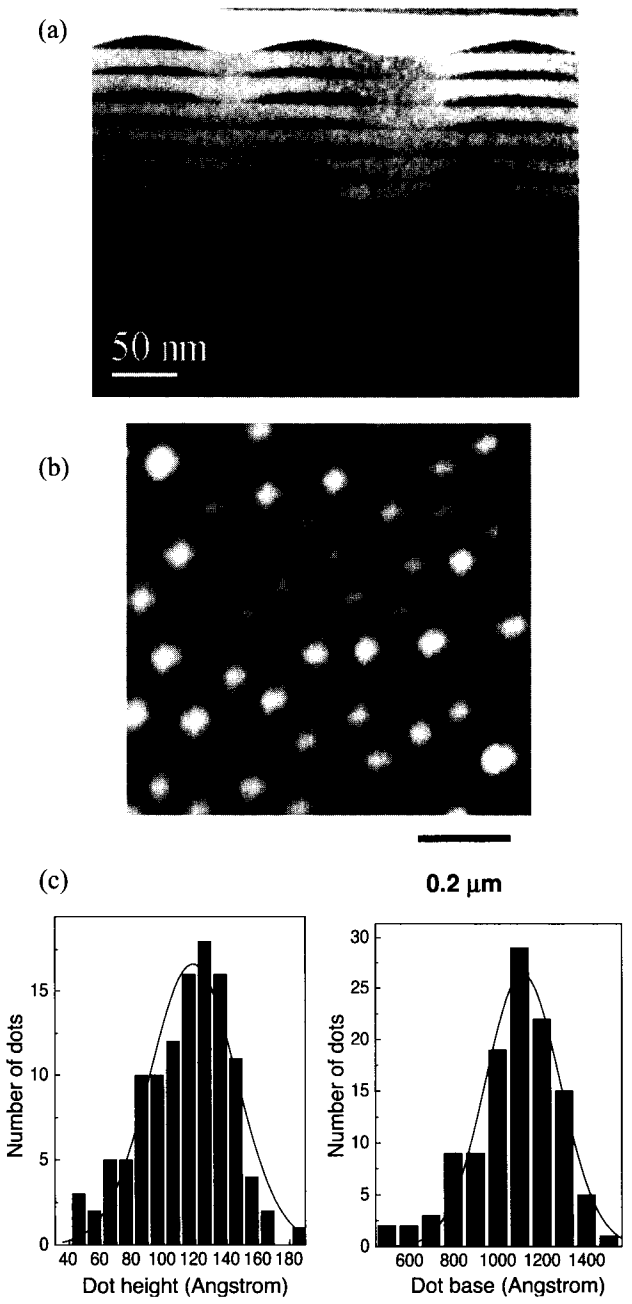


Fig. 1 (a) TEM image of sample A. Vertically correlated islands are evident. (b) AFM image of sample A. Most of the islands appear as square-based pyramids and the island density is determined to be 3.6×10^9 cm^{-2}. (c) Statistical base diameter and height distributions of the Ge islands of sample A. A typical Gaussian peak is used to fit the distribution data. The most probable height and base diameter are 11.9 and 110.4 nm, respectively.

18 Å, respectively. Samples D and E were grown with the same Ge layer thickness of 15 Å but at different growth temperatures of 500°C and 600°C, respectively. In addition, sample D has 10-period Ge and Si bilayers while sample E contains 22 periods. The Si spacer layer thickness of 20 nm was used for all superlattices. No cap Si layers were grown intentionally on these samples for convenience of atomic force microscopy (AFM) characterizations. Transmission electron microscopy measurements were exploited as well to obtain the size, density, and uniformity of the self-assembled dots. Raman scattering measurements were performed on a Renishaw Raman Imaging 2000 system with a 514 Ar^+ laser as the excitation light source in the backscattering configuration.

Fig. 1a shows a typical cross-sectional TEM image of sample A. Ten-period vertically correlated layers of Ge dots are evident. The dots in the different layers (beyond the third Ge layer) are identical in size and shape. This observation is also seen for other dot samples. Fig. 1b and c shows an AFM image, and height and base statistic analysis of sample A, respectively. Gaussian distributions were used to fit the data. Most of the dots appear as pyramid and the average dot base and height are 11.9 and 110.4 nm, respectively. Similar measurements have been performed on other samples as well. The growth parameters and structural data were summarized in Table 1. It should be noted that the structural data were from AFM measurements only and not calibrated by TEM characterizations.

Optical Phonons

Fig. 2 shows the Raman spectra of the dot samples in the spectral region of SiGe and GeGe optical phonons. Similar SiGe (from 400 to 420 cm^{-1}) and GeGe lines (near 300 cm^{-1}) are observed for dot samples. The frequency position of SiGe optical phonons more or less represents the degree of the interdiffusion between Si spacers and Ge dots. Samples A, B, and C have almost the same Si–Ge optical phonon line shape. The peak frequency changes slightly from 418 cm^{-1} for sample A with a nominal Ge thickness of 1.2 nm to 416 cm^{-1} for sample C with a nominal Ge thickness of 1.8 nm. To understand this, we shall notice that the interdiffusion arises from both thermal diffusion and strain-induced diffusion. As the same growth temperature of 540°C was used for the three samples, the degree of interdiffusion due to thermal management is the same. The slight difference in frequency therefore comes from the difference in strain-induced diffusion. The thicker Ge was deposited the more relaxation was presented in the dots (as will be discussed in detail in the following). Strain-induced interdiffusion in the samples with thicker nominal Ge thickness is smaller. The resulting lower frequency can be understood with the

Table 1 Growth parameters and structural data of the samples

Sample	Ge layer thickness (Å)	Si layer thickness (nm)	Growth T (°C)	Dot base (nm)	Dot height (nm)	Density (cm^{-2})
A	12	20	540	110.4	11.9	3.6×10^9
B	15	20	540	122.0	14.0	4.1×10^9
C	18	20	540	122.2	16.0	3.5×10^9
D	15	20	500	114.7	15.1	5.9×10^8
E	15	20	600	175.5	10.2	2.6×10^8

following picture: The eigen-frequency of these optical phonons is determined by the effective stretching force constant for each bond and the mass of atoms at the end of each bond. For example, SiSi optical phonon band is centered at 520 cm^{-1} for all of the samples and the substrate (not shown in the figure), considering the force constant of SiSi to be 39.5 N/m, and atomic weight of 28.1 for Si.[46] Likewise, optical phonons of crystal Ge can be found at 300 cm^{-1} with the GeGe bond force constant of 35 N/m and atomic weight of 72.6. Any mixture of Si and Ge leads to SiGe optical phonons with the frequency in between 300 and 520 cm^{-1}. For simplicity, we assume that Ge quantum dots are fixed and only Si atoms move. At a situation of stronger interdiffusion, more Si atoms are introduced into Ge dots. Thus SiGe optical phonon band has a frequency closer to SiSi optical phonon band, which is exactly the case for samples A through C. Likewise, much lower frequency position of 406 cm^{-1} for sample D and slightly larger frequency of about 420 cm^{-1} for sample E compared with that of sample B (due to the same nominal Ge) can also be explained by the above picture.

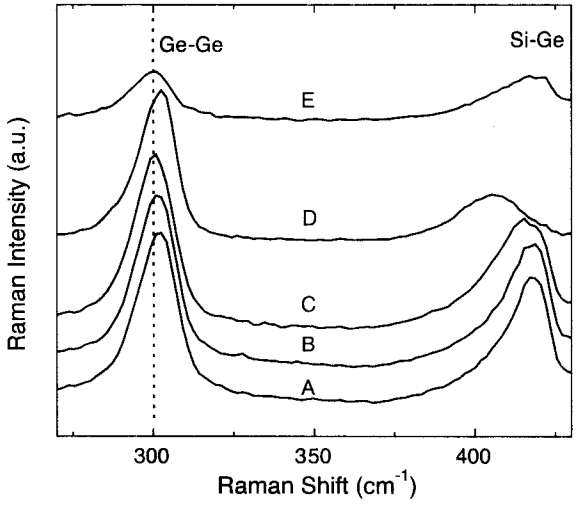


Fig. 2 Raman spectra of the dot samples. Similar SiGe and GeGe lines are observed for these samples. The frequency positions of the GeGe optical phonons in dot samples are shifted slightly to higher frequencies with respect to their bulk value (300 cm^{-1}).

Now let us analyze the GeGe optical phonons in the samples. It is important to note that the GeGe modes from the dot samples are different from the second-order transverse acoustic (2TA) phonon mode[47] for Si at 303 cm^{-1}. For comparison, the vertical dotted line is plotted at 300 cm^{-1} to represent the optical phonon position for bulk crystalline Ge. The frequency positions of the GeGe optical phonons in dot samples are shifted slightly to higher frequencies with respect to their bulk value. There are several mechanisms to cause a Raman shift of GeGe optical phonons. The first one is phonon confinement. It is known that the optical phonon branches of bulk Ge are quadratic and, moreover, nearly flat at the Brillouin zone center ($k \cong 0$). Confined optical phonons in a nanocrystal are equivalent to those vibrations in an infinite crystal whose wave vector is given by $m\pi/d$, where m is an integer and d is the size of the nanocrystal, in our case, the height of the dots. In the dots with a very small height of 1.5–2 nm,[21] the phonon confinement effect was observed to result in a shift less than 2 cm^{-1}. As the dot size for the present samples is much larger than the lattice constant of Ge, the wave vector is extremely small, leading to insignificant phonon confinement effect.

The size confinement effect of optical phonons in quantum dots with small size should give rise to a shift of phonon band to lower frequency side due to the negative dispersion of the optical phonon branch. A compressive strain on the dots in the lateral directions as a result of the lattice mismatch of Si and Ge, however, leads to a GeGe mode shift to the higher frequency side. This concept can be written as:[48]

$$\omega = \omega_0 + \frac{1}{2\omega_0}[p\varepsilon_{zz} + q(\varepsilon_{xx} + \varepsilon_{yy})] \quad (1)$$

where ω is the Ge-Ge mode frequency induced by a biaxial strain, ω_0 is the frequency of the Ge zone-center LO phonon; p and q are the Ge deformation potentials; ε_{xx} and ε_{yy} are the biaxial strain $(a-a_0)/a$ with a and a_0 being unstrained and strained lattice constants, respectively. The physical parameters used in Eq. 1 can be obtained from

Table 2 The experimental GeGe mode frequencies, integrated peak intensity ratios, and calculated Ge composition in dots and residual in-plane strain in the dot samples

Sample	ω(GeGe) (cm^{-1})	ω(SiGe) (cm^{-1})	I_{GeGe}/I_{SiGe}	Ge composition in dots	ε_{xx} (10^{-3})
A	302	418	1.45	0.48	−4.8
B	301.1	417.5	1.45	0.48	−2.65
C	300.5	416	1.45	0.48	−1.2
D	302.4	406	2.14	0.58	−5.8
E	300.4	419.5	1.29	0.45	−0.96

Ref. [34]. For fully strained pure Ge on Si, we obtain ω=317.4 cm^{-1}. The large difference between this calculated number and the experimental GeGe mode values (close to 300 cm^{-1} as shown in Fig. 2) suggests that the dots are not fully strained as assumed in the calculation. By using Eq. 1 and the experimental GeGe optical phonon frequencies ω, the residual biaxial strain (ε_{xx}) on the quantum dots can be estimated and listed in Table 2. Very small residual strain values obtained suggest the dots are almost fully relaxed.

An obvious reason of the strain relaxation is the atomic intermixing at the Si–Ge interface. The degree of the interface intermixing can be determined by the integrated peak intensity ratio I_{GeGe}/I_{SiGe} as the intensity depends on the relative number of corresponding bonds. This concept was used for SiGe alloy[49] and Ge quantum dots[37] and can be expressed as:

$$\frac{I_{GeGe}}{I_{SiGe}} \cong B\frac{x}{2(1-x)} \tag{2}$$

where x is the average Ge concentration and the coefficient B is related to the Bose factor and the frequencies of GeGe and SiGe optical modes of the alloy. It is found that the coefficient B varies weakly with alloy composition and is determined to be 3.2.[50] I_{GeGe}/I_{SiGe} can be obtained from the spectra in Fig. 2. The Ge concentrations in quantum dots estimated by this method are listed in Table 2. It should be noted that the estimation by such interpolation is rough because the Si 2TA phonons and optical phonons from wetting layers have not been subtracted in the calculation. Moreover, the composition obtained is an average value, which does not show any information about composition distribution in the dots.

Fig. 3 plots the GeGe peak position as a function of the Ge concentration in the dots. Symbols are experimental data. Solid straight lines 1 and 2 are used to guide the eyes and indicate that the different GeGe mode frequency is obtained for dots not only with different Ge concentrations but also with a similar Ge concentration. Solid line 3 is a calculation result using Eq. 1. In this calculation, we assume the same ω_0 and other deformation potentials for pure Ge. This approximation is reasonable because the interdiffusion occurs in the interface and the "core" of the dots remains pure Ge. For example, the nonuniform Ge content in the dots has been proved by elemental distribution analysis using an electron energy filtering imaging method in TEM[51] and scanning TEM.[45] The in-plane strain ε_{xx} used can be written as $(a_{Si_{1-x}Ge_x}-a_0)/a_{Si_{1-x}Ge_x}$ with the lattice parameter of an alloy with Ge concentration of x determined by Vegard's law $a_{Si_{1-x}Ge_x}=xa_{Ge}+(1-x)a_{Si}$, where a_{Ge} and a_{Si} are lattice parameters of pure Ge and Si, respectively. Simple calculations show, for example, in-plane strain to be −0.02 for an alloy with Ge concentration of 0.48 (samples A, B, and C), leading to a GeGe mode frequency of 308.3 cm^{-1}. The calculated values show a considerable shift from the value obtained for fully strained Ge on Si (317.4 cm^{-1}); however, these are still larger than the peak frequencies of the GeGe modes of the samples A through E. As a matter of fact, the absolute value of the residual strain on the dots shown in Table 2 is much smaller than the absolute value estimated here by interdiffusion only.

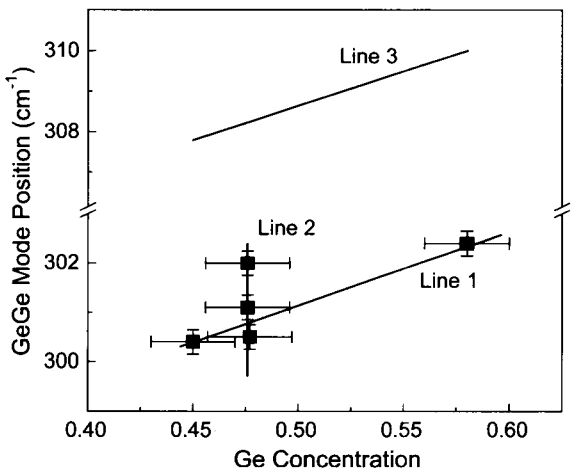


Fig. 3 GeGe mode position as a function of Ge concentration in dots. Ge concentration is obtained by means of the integrated intensity ratio of optical phonons of GeGe and SiGe. Solid lines 1 and 2 are used to guide the eyes. The results show that the different GeGe mode frequency is obtained for dots not only with different Ge concentrations but also with a similar Ge concentration. Line 3 is the calculated data by assuming Si/Ge interdiffusion only.

Moreover, the picture of interdiffusion does not completely account for the peak position difference for the different samples with the same Ge composition as of line 3 in Fig. 3. The above analysis suggests that additional strain relaxation mechanisms exist apart from Ge/Si interdiffusion. We believe that this additional strain relaxation comes from the process of the formation of Ge dots.

The relaxation of the dots in addition to atomic interdiffusion consists of the relaxation of the surface-layered dots and the relaxation of the embedded dots. The relaxation of the surface-layered dots is related to the nature of Stranski–Krastanov growth mode. When Ge is deposited on Si, it first takes a layer-by-layer growth format. As the Ge film exceeds its critical thickness, it becomes rough by forming pyramidal islands to relieve strain energy. As more Ge is deposited, the islands transform from the pyramid shape to the dome shape, leading to more strain relaxation. Atomic force microscopy measurements on our samples indeed showed, for example, the pyramid-over-dome ratio decreases as the increase of the Ge deposition for these samples grown at the same temperature of 540°C.[25] The relaxation is also induced by the dot embedding effect. Sutter and Lagally have used a strain-driven atom diffusion model to show that the dot shape transforms to truncated shape after the growth of Si layer on top.[52] Indeed, we have observed this transition from the cross-section TEM measurements. Because of the capping process, the in-plane strain is relaxed.[39]

After recognizing all the above relaxation mechanisms, we can fully explain the observed GeGe optical phonon frequencies. First, let us take a look at line 1 of Fig. 3. This line represents the GeGe mode trend of a set of samples (C, D, and E) grown at different temperatures. GeGe mode frequency changes from 302.4 to 300.4 cm^{-1} when the Ge concentration decreases from 0.58 to 0.45 as a result of growth temperature from 500°C to 600°C, indicating that the sample grown at higher temperature induces stronger interface intermixing, or more strain relaxation. Quantitative calculations using Eq. 1 show that a shift of 2.2 cm^{-1} is obtained for the present interdiffusion difference in the three samples (as can be seen from line 3 in Fig. 3). This number is almost the same as the experimental observed shift, suggesting that the strain relaxation from the different degrees of interdiffusion is the main factor responsible for the different frequencies of the GeGe mode in this set of samples. In other words, the overall effect of the dot shape transition on the GeGe mode frequency is about the same for the three samples. Line 2 in Fig. 3 represents another set of samples (A, B, and C). As shown previously, this set of samples has the same Ge concentration in the dots. Nevertheless, the GeGe optical phonon frequency is different and the largest difference is 1.5 cm^{-1}. This suggests different strains are present on the dots in different samples, which arises from the different dot morphologies in the different samples, such as pyramid-over-dome ratios, dot height-over-base ratios, and embedding effect.

The above-analyzed, negligible phonon confinement effect but dominant atomic intermixing and/or strain effect in determining frequencies of optical phonons for quantum dot samples can also be proved by performing annealing experiments. Fig. 4a shows the Raman spectra of sample A under rapid thermal annealing (RTA) for 5 min at different temperatures. The vertical dotted line is plotted at the center of GeGe mode from as-grown sample for comparison. The GeGe Raman line broadens and shifts to lower frequencies as the annealing temperature increases. This phenomenon was also observed when annealing other samples. Fig. 4b shows the annealing temperature dependence of the Raman shift for the dot samples grown at different temperature. After 850°C annealing for 5 min, the GeGe optical mode shifts downward by as much as 8, 6–7, and 5 cm^{-1} for the samples grown at 500°C, 540°C, and 600°C, respectively. The intermixing after annealing leads to fully relaxed SiGe alloy-like dots, causing the observed peak shift.

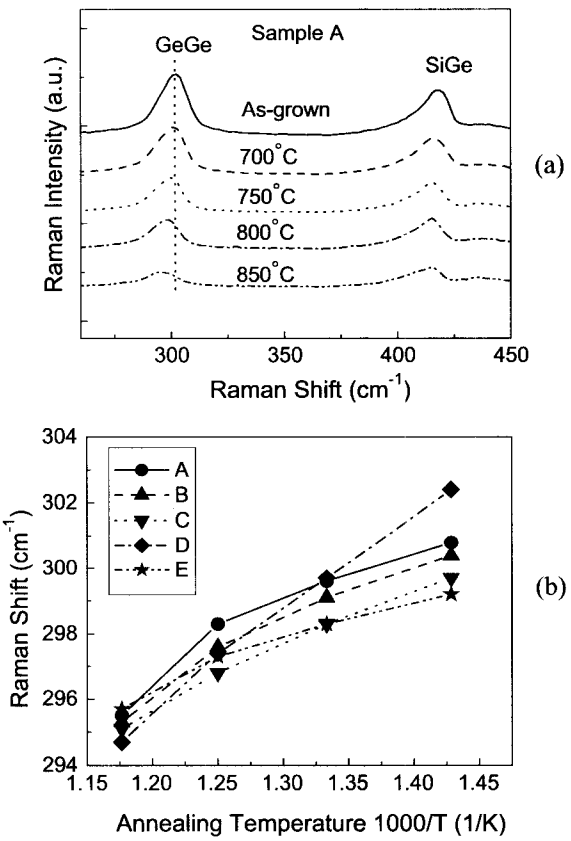


Fig. 4 (a) Raman spectra of sample A under RTA for 5 min at different temperatures, and (b) Raman shift as a function of annealing temperature for different samples.

Acoustic Phonons

Fig. 5a shows the Raman spectra of the samples A, B, and C, which were grown with the same temperature and different Ge layer thickness. The vertical dot lines are used to facilitate the comparison among the samples. Similar SiSi, SiGe, and GeGe optical vibration peaks are seen at around 520, 415, and 300 cm^{-1}, respectively. Besides these peaks, low-frequency oscillation peaks are observed for all of the three samples around 84.3, 104.7, 126.6, 150, and 171.8 cm^{-1}, respectively. These nearly equal-distance peaks with a period of about 20 cm^{-1} are higher-order acoustic phonons. The lower orders were attenuated by the Rayleigh line rejection filter and thus were not observed. An interesting phenomenon is that the acoustic peak frequencies do not scale with the dot size. Lamb's theory predicts that confined phonon frequency should be inversely proportional to the nano-crystal size[53] and was shown to be true for a number of nano-crystal systems such as CdS_xSe_{1-x} in glasses[54] and Si in SiO_2 matrix.[55] The present result suggests that the observed low-frequency peaks are not confined phonon modes in quantum dots.

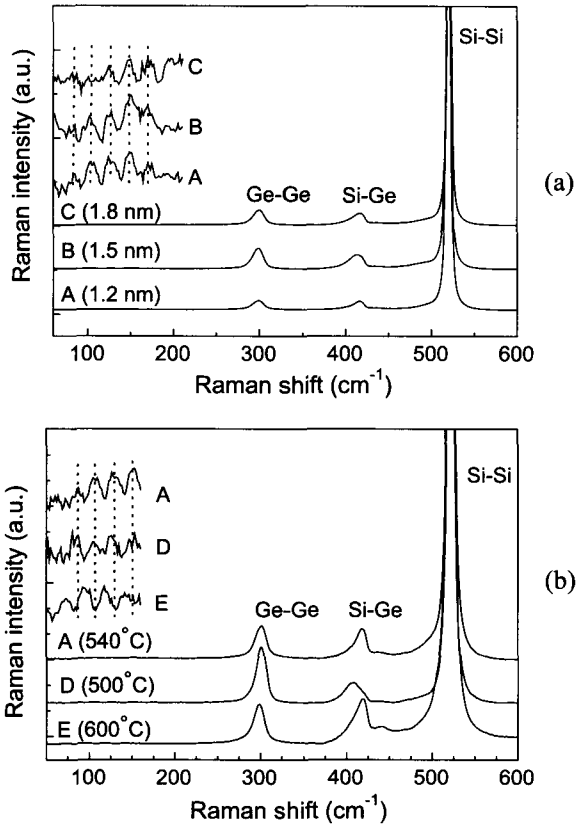


Fig. 5 Raman spectra of the samples with (a) different nominal Ge thicknesses of 1.2, 1.5, and 1.8 nm for samples A, B, and C, respectively, and (b) with different growth temperatures of 540°C, 500°C, and 600°C for samples A, D, and E, respectively. Acoustic phonon modes with the same frequencies are observed for different Ge-thickness samples, while different frequencies and periods are observed for different samples with different growth temperatures.

Fig. 5b shows the Raman spectra of the samples A, D, and E, which were grown at different temperatures. The vertical dot lines are again used here to facilitate the comparison among the samples. Acoustic phonons for different samples have different frequencies. Specifically, acoustic phonon peak frequencies for sample A have been assigned in Fig. 5a. Sample D has acoustic phonons around 71.4, 95.7, 118.8, and 141 cm^{-1}, and sample E has acoustic phonons around 83.8, 105.8, 126.8, and 149.8 cm^{-1}. The period of these nearly equal-distance peaks is about 24, 22, and 20 for samples D, E, and A, respectively.

There are two possibilities, which may be used to explain the observed periodic acoustic phonons.[56] The first one is that these are probably phonon modes confined in quantum wells where zone-edge phonon modes have been made, "allowed" by breakdown of wave vector conservation. The second one is that they are associated with a superlattice where low-frequency zone-edge phonon modes have been folded into zone center. The first concept has been used to explain the periodic acoustic phonon oscillations in the InAs/InP self-assembled quantum dots.[12] The period of the periodic acoustic phonons can be expressed by:

$$\Delta\omega = \frac{1}{2L} \cdot \frac{v}{c} \tag{3}$$

where L is the thickness of the quantum layer, v is the sound velocity of the phonons, and c is the speed of the light. If the observed phonons were the confined phonons in the Si spacer layers, then with the velocity $v_{Si} = 8.44 \times 10^5$ cm/sec[57] and the peak periods of 20–24 cm^{-1} for the present samples, we would obtain the Si layer thicknesses for our samples to be around 5.9–7 nm. The designed thickness of Si spacer thickness for all the samples was 19–20 nm. The actual effective thickness of Si spacer layers has shrunk to around 14 nm characterized by cross-sectional TEM. The reason is that Si spacers were deposited on curvature surfaces containing three-dimensional dots. A similar situation was often observed when growing short-period superlattices on relaxed virtue substrates, which consisted of cross-hatch undulation patterns. The layer deposited on top was stretched as a result of undulation and the effective thickness along the growth direction was shrunk. The effective thickness of about 14 nm for the present samples, however, is still more than two times of the calculated 5.9–7 nm, suggesting that these phonons have nothing to do with the confinement in the Si spacers. Likewise, the phonons are not confined phonons in the wetting layers either. Samples grown at different temperatures lead to different

wetting layer thickness because of Si/Ge interdiffusion. Different from the dots grown at much lower temperature (300°C)[26] or by the use of antimony as surfactant,[34] where the wetting layers were more or less around 3–4 MLs, the wetting layers in this set of samples had the thickness in the region of 15–20 Å with cross-sectional TEM characterizations (not shown here). Therefore the wetting layers are not pure Ge but SiGe alloy. Phonon velocity for SiGe can be obtained by the linear interpolation of LA phonon velocity of Si and Ge. Simple calculations showed that one could never obtain an agreeable solution compared with the experimental results.

The above analysis leaves the only possible origin for these phonons, i.e., they are folded acoustic modes related to the periodicity of the superlattice. Rytov's elastic continuum model[58] has been well applied for folded acoustic phonons in quantum well superlattices and was also seen to explain the acoustic phonon features in Ge dot multilayers[26–29] and InAs/GaAs dot multilayers.[14] In this model, the acoustic phonon dispersion can be written as:

$$\cos(qd) = \cos\left(\frac{\omega d_1}{V_1}\right)\cos\left(\frac{\omega d_2}{V_2}\right) - \frac{1}{2}\left(R+\frac{1}{R}\right) \times \sin\left(\frac{\omega d_1}{V_1}\right)\sin\left(\frac{\omega d_2}{V_2}\right) \quad (4)$$

where q is the superlattice wave vector perpendicular to the layers and is roughly determined to be $q \cong 4\pi n'/\lambda_L$, where λ_L is the incident laser light wavelength and n' is the refractive index of the material at that wavelength. In addition, $R=(V_1\rho_1)/V_2\rho_2$, $d=d_1+d_2$. d_1 and d_2, ρ_1 and ρ_2, and V_1 and V_2 are the thickness, density, and sound velocity in Ge and Si layers, respectively. The physical parameters can be obtained from Ref. [57]. The model designed for quantum well superlattice can be considered as a zero-order approximation to the real quantum dot superlattice as it does not take into account quantum dot shape, size, and density. It was suggested in Ref. [27] to use nominal thickness of Ge layer as an effective thickness d_2. Nevertheless, with this rough approximation, good agreement with experimental data was achieved.[26–29] Following this approach, with the effective Si spacer layer thicknesses of about 14 nm and wetting layer thicknesses of about 1.5 nm in our samples, we calculated the phonon dispersion and showed acoustic phonon peak period of 19–21 cm^{-1}, which was close to the experimentally observed 20–24 cm^{-1}. Therefore we conclude that the observed low-frequency modes are folded acoustic phonons. More accurate description of the acoustic spectra of semiconductor quantum dot superlattices will require rigorous calculations of acoustic phonon scattering on quantum dot as well as multiple scattering effect produced by quantum dot ensemble in each layer.

CONCLUSION

We have studied Raman scattering from optical phonons and acoustic phonons in the Ge quantum dot superlattices. The GeGe optical phonon frequencies of the dot samples were larger than 300 cm^{-1} but within 3 cm^{-1}, indicating that the dots were relaxed in the superlattices. Strain relaxation was found not only from the atomic interdiffusion but also from the dot morphology transition both for the surface dots and the embedded dots. The annealing experiments also proved that the strain relaxation and intermixing were the key parameters for the observed optical phonons. Optical phonon studies show that Raman spectroscopy is an efficient tool for the determination of the average compositions and strains for zero-dimensional quantum dots. Periodic oscillations in the lower frequency region were observed for quantum dot superlattice samples. These acoustic phonons were found to be related to superlattice period and could be well explained by the elastic continuum model.

Furthermore, the method for analysis of phonons in the above Ge quantum dot superlattice system can be generally used for other quantum dot systems, such as, InGaSb/GaAs, InSb/InP, InGaAs/GaAs, InAs/InP, AlGaAs/InAs, InAs/AlAs, and GaN/(Si)AlGaN, CdSe/ZnSe, and so on. Phonon spectra in these material systems depend on the physical parameters of superlattice, such as concentration in dots, dot size, and period thickness. In the case of very small dots with high density (e.g., greater than 10^{11} cm^{-2}), multiple scattering effects may play an important role in the modification of phonon spectra. The phonon densities of states in quantum dot superlattices change as seen from the strong modification of phonon spectra with the appearance of quantum dots and can be used to calculate the thermal conductivities of superlattices. Research has shown a great reduction of thermal conductivities in quantum dot superlattices, which have great applications in the fabrication of new-generation solid-state thermoelectric devices. Phonon spectra modification by quantum dots also suggests the potential applications of quantum dot superlattices as phonon lenses, filters, and reflectors for phonon-assisted lasers in indirect-band gap material systems.

ACKNOWLEDGMENTS

The authors wished to thank Dr. S. G. Thomas of Motorola Semiconductor Products Sector for performing some of the TEM characterizations and Prof. Peter Y. Yu of UC Berkeley for valuable discussions. This work was funded by the DoD MURI-ONR program on Thermoelectrics (N00014-97-1-0516) and the AFOSR MURI program on phonon engineering (F49620-00-1-0328).

REFERENCES

1. Takagahara, T.; Takeda, K. Theory of the quantum confinement effect on excitons in quantum dots of indirect-gap materials. Phys. Rev., B **1992**, *46*, 15578.
2. Ren, S.Y. Quantum confinement of edge states in Si crystallites. Phys. Rev., B **1997**, *55*, 4665.
3. Bennett, B.R.; Shanabrook, B.V.; Magno, R. Phonons in self-assembled (In,Ga,Al)Sb quantum dots. Appl. Phys. Lett. **1996**, *68*, 958.
4. Armelles, G.; Utzmeier, T.; Postigo, P.A.; Briones, F.; Ferrer, J.C.; Peiro, P.; Cornet, A. Raman scattering of InSb quantum dots grown on InP substrates. J. Appl. Phys. **1997**, *81*, 6339.
5. Zanelatto, G.; Pusep, Yu.A.; Moshegov, N.T.; Toropov, A.I.; Basmaji, P.; Galzerani, J.C. Raman study of the topology of InAs/GaAs self-assembled quantum dots. J. Appl. Phys. **1999**, *86*, 4387.
6. Artus, L.; Cusco, R.; Hernandez, S.; Patane, A.; Polimeni, A.; Henini, M.; Eaves, L. Quantum-dot phonons in self-assembled InAs/GaAs quantum dots: Dependence on the coverage thickness. Appl. Phys. Lett. **2000**, *77*, 3556.
7. Chu, L.; Zrenner, A.; Bichler, M.; Bohm, G.; Abstreiter, G. Raman spectroscopy of In(Ga)As/GaAs quantum dots. Appl. Phys. Lett. **2000**, *77*, 3944.
8. Toda, Y.; Moriwaki, O.; Nishioka, M.; Arakawa, Y. Resonant Raman scattering of optical phonons in self-assembled quantum dots. Physica, E **2000**, *8*, 328.
9. Galzerani, J.C.; Puseb, J.A. Raman spectroscopy characterization of InAs self-assembled quantum dots. Physica, B **2002**, *316*, 455.
10. Groenen, J.; Priester, C.; Carles, R. Strain distribution and optical phonons in InAs/InP self-assembled quantum dots. Phys. Rev., B **1999**, *60*, 16013.
11. Prieto, J.A.; Armelles, G.; Groenen, J.; Carles, R. Size and strain effects in the E1-like optical transitions of InAs/InP self-assembled quantum dot structures. Appl. Phys. Lett. **1999**, *74*, 99.
12. Huntzinger, J.R.; Groenen, J.; Cazayous, M.; Mlayah, A.; Bertrn, N.; Paranthoen, C.; Dehaese, O.; Carrere, H.; Bedel, E.; Armelles, G. Acoustic-phonon Raman scattering in InAs/InP self-assembled quantum dots. Phys. Rev., B **2000**, *61*, R10547.
13. Yin, J.; Wang, X.; Yin, Z.; Du, G.; Yang, S. InAs self-assembled quantum dots on GaAs/InP by low-pressure metal-organic chemical vapour deposition. Semicond. Sci. Technol. **2001**, *16*, 715.
14. Tenne, D.A.; Haisler, V.A.; Toropov, A.I.; Bakarov, A.K.; Gutakovsky, A.K.; Zahn, D.R.T; Shebanin, A.P. Raman study of self-assembled GaAs and AlAs islands embedded in InAs. Phys. Rev., B **2000**, *61*, 13785.
15. Tenne, D.A.; Haisler, V.A.; Bakarov, A.K.; Toropov, A.I.; Gutakovsky, A.K.; Shebanin, A.P.; Zahn, D.R.T. Self-assembled islands in the (Ga,Al)As/InAs heteroepitaxial system studied by Raman spectroscopy. Phys. Status Solidi, B **2001**, *224*, 25.
16. Tenne, D.A.; Bakarov, A.K.; Toropov, A.I.; Zahn, D.R.T. Raman study of self-assembled InAs quantum dots embedded in AlAs: Influence of growth temperature. Physica, E **2002**, *13*. 199.
17. Kuball, M.; Gleize, J.; Tanaka, S.; Aoyagi, Y. Resonant Raman scattering on self-assembled GaN quantum dots. Appl. Phys. Lett. **2001**, *78*, 987.
18. Kuball, M.; Gleize, J.; Tanaka, S.; Aoyagi, Y. On phonon confinement effects and free carrier concentration in GaN quantum dots. Phys. Status Solidi, B **2001**, *228*, 195.
19. Rho, H.; Smith, L.M.; Jackson, H.E.; Lee, S.; Dobrowolska, M.; Furdyna, J.K. Interface phonons in CdSe/ZnSe self-assembled quantum dot structures. Phys. Status Solidi, B **2001**, *224*, 165.
20. Liu, J.L.; Tang, Y.S.; Wang, K.L.; Radetic, T.; Gronsky, R. Raman scattering from a self-organized Ge dot superlattice. Appl. Phys. Lett. **1999**, *74*, 1863.
21. Liu, J.L.; Jin, G.; Tang, Y.S.; Luo, Y.H.; Wang, K.L.; Yu, D.P. Optical and acoustic phonon modes in self-organized Ge quantum dot superlattices. Appl. Phys. Lett. **2000**, *76*, 586.
22. Liu, J.L.; Tang, Y.S.; Wang, K.L. Response to "comment on 'Raman scattering from a self-organized Ge dot superlattices'". Appl. Phys. Lett. **1999**, *75*, 3574.
23. Liu, J.L.; Jin, G.; Tang, Y.S.; Luo, Y.H.; Wang, K.L.; Yu, D.P. Response to "Comment on 'Optical and acoustic phonon modes in self-assembled Ge quantum dot superlattices'". Appl. Phys. Lett. **2001**, *78*, 1162.
24. Liu, J.L.; Jin, G.; Tang, Y.S.; Luo, Y.H.; Lu, Y.; Wang, K.L.; Yu, D.P. Study of phonons in self-organized multiple Ge quantum dots. J. Electron. Mater. **2000**, *29*, 554.
25. Liu, J.L.; Wan, J.; Jiang, Z.M.; Khitun, A.; Wang, K.L.; Yu, D.P. Optical phonons in self-assembled Ge quantum dot superlattices: Strain relaxation effects. J. Appl. Phys. **2002**, *92*, 6804.
26. Milekhin, A.; Stepina, N.P.; Yakimov, A.I.; Nikiforov, A.I.; Schulze, S.; Zahn, D.R.T. Raman scattering of Ge dot superlattices. Eur. Phys. J., B **2000**, *16*, 355.
27. Milekhin, A.; Stepina, N.P.; Yakimov, A.I.; Nikiforov, A.I.; Schulze, S.; Zahn, D.R.T. Raman scattering studies of Ge dot superlattices. Appl. Surf. Sci. **2001**, *175*, 629.
28. Milekhin, A.; Nikiforov, A.I.; Pchelyakov, O.P.; Schulze, S.; Zahn, D.R.T. Phonons in self-assembled Ge/Si structures. Physica, E **2002**, *13*, 982.
29. Milekhin, A.; Nikiforov, A.I.; Pchelyakov, O.P.; Schulze, S.; Zahn, D.R.T. Self-selective Raman scattering in self-assembled Ge/Si quantum dot superlattices. Nanotechnology **2002**, *13*, 55.

30. Teo, K.L.; Qin, L.; Noordin, I.M.; Karunasiri, G.; Shen, Z.X.; Schmidt, O.G.; Eberl, K.; Queisser, H.J. Effects of hydrostatic pressure on Raman scattering in Ge quantum dots. Phys. Rev., B **2001**, *63*, 121306.
31. Qin, L.; Teo, K.L.; Shen, Z.X.; Peng, C.S.; Zhou, J.M. Raman scattering of Ge/Si dot superlattices under hydrostatic pressure. Phys. Rev., B **2001**, *64*, 075312.
32. Cazayous, M.; Huntzinger, J.R.; Groenen, J.; Mlayah, A.; Christiansen, S.; Strunk, H.P.; Schmidt, O.G.; Eberl, K. Resonant Raman scattering by acoustic phonons in Ge/Si self-assembled quantum dots: Interferences and ordering effects. Phys. Rev., B **2000**, *62*, 7243.
33. Cazayous, M.; Groenen, J.; Huntzinger, J.R.; Mlayah, A.; Schmidt, O.G. Spatial correlations and Raman scattering interferences in self-assembled quantum dot multilayers. Phys. Rev., B **2001**, *64*, 033306.
34. Kwok, S.H.; Yu, P.Y.; Tung, C.H.; Zhang, Y.H.; Li, M.F.; Peng, C.S.; Zhou, J.M. Confinement and electron-phonon interactions of the E_1 exciton in self-organized Ge quantum dots. Phys. Rev., B **1999**, *59*, 4980.
35. Persans, P.D.; Deelman, P.W.; Stokes, K.L.; Schowalter, L.J.; Byrne, A.; Thundat, T. Optical studies of Ge islanding on Si(111). Appl. Phys. Lett. **1997**, *70*, 472.
36. Guedj, C.; Beyer, A.; Muller, E.; Grutzmacher, D. Raman spectroscopy of carbon-induced germanium dots. Appl. Phys. Lett. **2001**, *78*, 1742.
37. Groenen, J.; Carles, R.; Christiansen, S.; Albrecht, M.; Dorsch, W.; Strunk, H.P.; Wawra, H.; Wagner, G. Phonons as probes in self-organized SiGe islands. Appl. Phys. Lett. **1997**, *71*, 3856.
38. Gaiduk, P.I.; Larsen, A.N.; Hansen, J.L.; Mudryj, A.V.; Samtsov, M.P.; Demenschenok, A.N. Self-assembly of epitaxially grown Ge/Si quantum dots enhanced by As ion implantation. Appl. Phys. Lett. **2001**, *79*, 4025.
39. Liu, N.; Tersoff, J.; Baklenov, O.; Holmes, A.L., Jr.; Shih, C.K. Nonuniform composition profile in $In_{0.5}Ga_{0.5}As$ alloy quantum dots. Phys. Rev. Lett. **2000**, *84*, 334.
40. Rosenauer, A.; Fischer, U.; Gerthsen, D.; Forster, A. Composition evaluation of $In_xGa_{1-x}As$ Stranski–Krastanow-island structures by strain state analysis. Appl. Phys. Lett. **1997**, *71*, 3868.
41. Walther, T.; Humphreys, C.J.; Cullis, A.G. Observation of vertical and lateral Ge segregation in thin undulating SiGe layers on Si by electron energy-loss spectroscopy. Appl. Phys. Lett. **1997**, *71*, 809.
42. Chaparro, S.A.; Drucker, J.; Zhang, Y.; Chandrasekhar, D.; McCartney, M.R.; Smith, D.J. Strain-driven alloying in Ge/Si(100) coherent islands. Rev. Lett. **1999**, *83*, 1199.
43. Wiebach, T.; Schmidbauer, M.; Hanke, M.; Raidt, H.; Kohler, R. Strain and composition in SiGe nanoscale islands studied by x-ray scattering. Phys. Rev., B **2000**, *61*, 5571.
44. Jiang, Z.M.; Jiang, X.M.; Jiang, W.R.; Jia, Q.J.; Zheng, W.L.; Qian, D.C. Lattice strains and composition of self-organized Ge dots grown on Si(001). Appl. Phys. Lett. **2000**, *76*, 3397.
45. Harvey, A.; Davock, H.; Dunbar, A.; Bangert, U.; Goodhew, P.J. Three-dimensional compositional analysis of quantum dots. J. Phys., D, Appl. Phys. **2001**, *34*, 636.
46. Alonso, M.I.; Winer, K. Raman spectra of c-$Si_{1-x}Ge_x$ alloys. Phys. Rev., B **1989**, *39*, 10056.
47. Uchinokura, K.; Sekine, T.; Matsuura, E. Critical-point analysis of the two-phonon Raman spectrum of silicon. J. Phys. Chem. Solids **1974**, *35*, 171.
48. Cerdeira, F.; Buchenauer, C.J.; Pollak, F.H.; Cardona, M. Stress-induced shifts of first-order Raman frequencies of diamond- and zinc-blende-type semiconductors. Phys. Rev., B **1972**, *5*, 580.
49. Renucci, M.A.; Renucci, J.B.; Cardona, M., III. In *Disordered Systems: Raman Scattering in Ge–Si Alloys*, Proceedings of the 2nd International Conference on Light Scattering in Solids, Flammarion: Paris, 1971; 326.
50. Mooney, P.M.; Dacol, F.H.; Tsang, J.C.; Chu, J.O. Raman scattering analysis of relaxed Ge_xSi_{1-x} alloy layers. Appl. Phys. Lett. **1993**, *62*, 2069.
51. Liao, X.Z.; Zou, J.; Cockayne, D.J.H.; Qin, J.; Jiang, Z.M.; Wang, X.; Leon, R. Strain relaxation by alloying effects in Ge islands grown on Si (100). Phys. Rev., B **1999**, *60*, 15605.
52. Sutter, P.; Lagally, M.G. Embedding of nanoscale 3D SiGe islands in a Si matrix. Phys. Rev. Lett. **1998**, *81*, 3471.
53. Lamb, H. Proc. Lond. Math. Soc. **1882**, *13*, 189.
54. Champagnon, B.; Andrianasolo, B.; Ramos, A.; Gandais, M.; Allais, M.; Benoit, J. Size of Cd(S,Se) quantum dots in glasses: Correlation between measurements by high-resolution transmission electron microscopy, small-angle x-ray scattering, and low-frequency inelastic Raman scattering. J. Appl. Phys. **1993**, *73*, 2775.
55. Fujii, M.; Kanzawa, Y.; Hayashi, S.; Yamamoto, K. Raman scattering from acoustic phonons confined in Si nanocrystals. Phys. Rev., B **1996**, *54*, 8373.
56. Yu, P.Y. Comment on "optical and acoustic phonon modes in self-organized Ge quantum dot superlattices". Appl. Phys. Lett. **2001**, *78*, 1160.
57. Lockwood, D.J.; Dharma-Wardana, M.W.C.; Baribeau, J.-M.; Houghton, D.C. Folded acoustic phonons in Si/Ge_xSi_{1-x} strained-layer superlattices. Phys. Rev., B **1987**, *35*, 2243.
58. Rytov, S.M. Akust. Z. **1956**, *2*, 71.

Quantum Dots, Self-Assembled: Calculation of Electronic Structures and Optical Properties

Andrew Williamson
Lawrence Livermore National Laboratory, Livermore, California, U.S.A.

INTRODUCTION

As evidenced by many of the articles in this encyclopedia, the past decade has witnessed a series of significant advances in the growth of semiconductor nanostructures. These nanostructures range in size from a few atoms up to several million atoms and have been produced by myriad synthesis techniques.[1] In this article, we describe calculations of the electronic and optical properties of self-assembled (so-called Stranski–Krastanow) quantum dots produced by molecular beam epitaxy.[2,3] Even more dramatic than the advances in the growth of these structures has been the development of extremely sophisticated techniques for measuring the optical, electronic, magnetic, and transport properties of these quantum dot systems. The availability of such high-quality measurements for a wide range of properties presents theorists with the formidable challenge of constructing models that can explain the origins of these experimental observations in terms of the underlying energy states of the quantum dots.

Calculating such energy states in self-assembled semiconductor quantum dots is rendered particularly challenging by a number of factors:

1. The quantum dots contain a large number of atoms. A typical self-assembled quantum dot has a base of $\sim$300 Å and a height of $\sim$50 Å. Therefore the dot itself may contain $\sim 10^5$ atoms. This dot then needs to be surrounded by a barrier material to isolate it from other dots. Therefore a representative system containing both the dot and barrier typically contains $\sim 10^6$ atoms.
2. By the nature of the growth process, self-assembled quantum dots are highly strained. For example, in the most common InAs/GaAs material combination, the lattice mismatch is 7%. Therefore an accurate solution of the strain profile in the system is required before the electronic structure can be calculated.
3. The valence band maximum (VBM) in III–V semiconductor materials is threefold degenerate (in the absence of strain and spin–orbit splitting), therefore any realistic approach must describe at least the band mixing between the three valence band edge states.
4. In InAs, the bulk band gap is 0.42 eV and the spin–orbit splitting is 0.38 eV, therefore the mixing between valence and conduction band states and split off states also have to be taken into account.
5. In a zero-dimensional InAs/GaAs quantum dot system, the charge carriers are artificially confined inside the dot, which is typically smaller than the bulk excitonic radius—the "Strong Confinement" regime. This dramatically enhances the Coulomb interaction between charges in the dot and strongly modifies the dielectric screening.

OVERVIEW

The standard theoretical approach to this problem has been to adapt effective mass-based techniques and their more sophisticated extension, the $\mathbf{k}\cdot\mathbf{p}$ method, which has been extremely successful in explaining a range of properties of bulk semiconductors. In essence, these methods expand the single-particle wavefunctions of the system in a basis of bulk Bloch orbitals derived from the Brillouin zone center (Γ point). If a sufficient number of basis states are included, this expansion provides an excellent description of the band structure of the bulk material close to the Γ point. However, errors in the predicted energy rapidly increase as one moves away from the zone center. To improve the description of the band structure away from the zone center, one typically includes more and more basis functions in the expansion of the wavefunctions. The successes of this effective mass model in describing spectroscopic and transport properties in both three-dimensional bulk systems and two-dimensional quantum well structures are well documented. This model was also demonstrated to be able to provide at least a qualitative picture of the energy states in zero-dimensional systems. However, some of this success is mitigated by the fact that often, the parameters in the model have to be refit to the nanostructure system itself. Recently, a direct comparison[4] between an eight-band $\mathbf{k}\cdot\mathbf{p}$ calculation and a full pseudopotential calculation for a pyramidal InAs quantum dot embedded within bulk GaAs showed a generally good agreement between the

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013612

two techniques. However, the higher symmetry imposed by the $\mathbf{k}\cdot\mathbf{p}$ approach acts to omit certain energy level splittings and polarization anisotropies. The $\mathbf{k}\cdot\mathbf{p}$ approach also overconfined both electrons and holes level by $\sim$50 meV.

Instead, in this article, we describe our method of choice, the empirical pseudopotential method (EPM) approach to calculating the energy states in semiconductor quantum dots. This approach has several advantages over the conventional effective mass approach to the problem:

1. Once a pseudopotential has been developed to describe the bulk system, there are no adjustable parameters for describing a heterostructure system.
2. The EPM method has the same accuracy, whether it is describing a three-dimensional bulk system, a two-dimensional quantum well, a one-dimensional quantum wire, or a zero-dimensional quantum dot.
3. The EPM method provides an *atomistic* description of the system. Therefore the correct symmetry of the underlying crystal lattice is reproduced. In addition, an atomistic description of interfaces between materials is also maintained.

We begin with a detailed description of each of the stages required in a typical EPM calculation of the energy states in a semiconductor quantum dot. Then we describe a series of applications of the EPM technique to studying a range of optical and electronic properties of self-assembled quantum dots.

DESCRIPTION OF PSEUDOPOTENTIAL TECHNIQUES

To calculate the energy states associated with various electronic excitations in self-assembled quantum dots requires three stages of calculation:

(1) Assume the shape and composition and compute the strain:
We first construct a supercell containing both the quantum dot and surrounding barrier material. Sufficient barrier material is used, so that when periodic boundary conditions are applied to the system, the electronic and strain interactions between dots in neighboring cells is negligible. The atomic positions within the supercell are relaxed by minimizing the strain energy described by an atomistic force field[5,6] including bond bending, bond stretching, and bond bending–bond stretching interactions. More details of the atomistic relaxation are given in the section "Calculation of the Strain Profile."

(2) Set up the pseudopotential single-particle equation:
A single-particle Schrödinger equation is set up at the relaxed atomic positions, $\{\mathbf{R}_{n\alpha}\}$

$$\hat{H}\psi_i(\mathbf{r}) = \left\{-\frac{\beta}{2}\nabla^2 + \sum_{n\alpha}\hat{v}_\alpha(\mathbf{r}-\mathbf{R}_{n\alpha})\right\}\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \quad (1)$$

The potential for the system is written as a sum of strain-dependent, screened atomic pseudopotentials, v_α, that are fit to bulk properties extracted from experiment and first-principles calculations For more details of the constructing of the Hamiltonian see the section "Constructing the Single-Particle Hamiltonian."

(3) Calculate the screened, interparticle many-body interactions:
The calculated single-particle wavefunctions are used to compute the electron–electron, electron–hole, and hole–hole direct, J_{ee}, J_{eh}, J_{hh}, and exchange K_{ee}, K_{eh}, K_{hh} Coulomb energies. For more details, see the section "Calculation of 'Two-Body' Interactions."

The main approximations involved in our method are as follows: 1) the fit of the pseudopotential to the experimental data of bulk materials is never perfect; and 2) we neglect self-consistent iterations in that we assume that the screened pseudopotential drawn from a bulk calculation is appropriate for the dot. Our numerical convergence parameters are: 1) the size of the GaAs barrier separating periodic images of the dots; and 2) the number of basis functions used in the expansion of the wavefunctions (see the sections "Expansion in a Plane Wave Basis" and "Expansion in a Linear Combination of Bloch Bands" for more details).

Calculation of the Strain Profile

To obtain at *atomistic* description of the strain profile in a heterostructure system, we construct an expression for the strain energy in terms of few-body potentials between actual atoms

$$E_{\text{strain}} = \sum_{ij} V_2(\mathbf{R}_i - \mathbf{R}_j) + \sum_{ijk} V_3(\hat{\Theta}_{ijk}) + \cdots \quad (2)$$

where V_2 is a two-body term, V_3 is a three-body function of the bond angle, $\hat{\Theta}_{ijk}$. The functional form of these terms is taken to be strain-independent. The strain is determined by minimizing E_{strain} with respect to atomic positions $\{\mathbf{R}\}$.

Our chosen expression for the elastic strain energy, is a generalization (G-VFF) of the original valence force field (VFF)[5,7] model. Our implementation of the VFF includes bond stretching, bond angle bending, and bond length/bond angle interaction terms. This enables us to accurately reproduce the C_{11}, C_{12}, and C_{44} elastic constants in

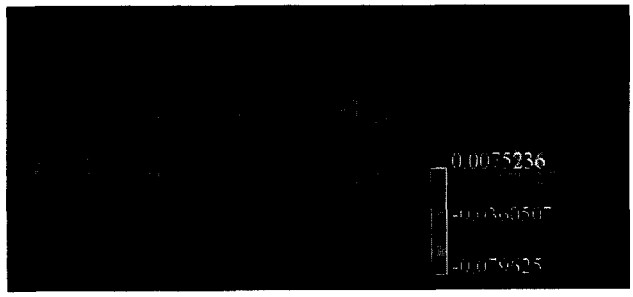


Fig. 1 Contour plot of the hydrostatic strain profile in a lens-shaped, self-assembled InAs/GaAs quantum dot with a base of 252 Å and a height of 35 Å. *(View this art in color at www.dekker.com.)*

a zincblende bulk material. We have also included higher-order bond stretching terms, which yield the correct dependence of the Young's modulus with pressure. The expression for the G-VFF total energy is:

$$E_{\text{G-VFF}} = \sum_i \sum_j^{nn_i} \frac{3}{8}[\alpha_{ij}^{(1)}\Delta d_{ij}^2 + \alpha_{ij}^{(2)}\Delta d_{ij}^3] + \sum_i \sum_{k>j}^{nn_i} \frac{3\beta_{jik}}{8d_{ij}^0 d_{ik}^0}[(\mathbf{R}_j - \mathbf{R}_i)\cdot(\mathbf{R}_k - \mathbf{R}_i) - \cos\ \theta_{jik}^0 d_{ij}^0 d_{ik}^0]^2 + \sum_i \sum_{k>j}^{nn_i} \frac{3\sigma_{ijk}}{d_{ik}^0}\Delta d_{ij} \times [(\mathbf{R}_j - \mathbf{R}_i)\cdot(\mathbf{R}_k - \mathbf{R}_i) - \cos\ \theta_{jik}^0 d_{ij}^0 d_{ik}^0] \quad (3)$$

where $\Delta d_{ij}=[|(R_i-R_j)^2 - d_{ij}^{02}|/d_{ij}^0]$. Here, $\mathbf{R}_i$ is the coordinate of atom I and d_{ij}^0 is the ideal (unrelaxed) bond distance between atom types of i and j. Also, θ_{jik}^0 is the ideal (unrelaxed) angle of the bond angle j–i–k. $\sum^{nn_i}$ denotes the summation over the nearest neighbors of atom i. Example values of α, β, and σ can be found in Ref. [8].

For an InGaAs alloy system, the bond angle and bond length/bond angle interaction parameters β, σ for the mixed cation Ga–As–In bond angle are taken as the algebraic average of the In–As–In and Ga–As–Ga values. In Fig. 1, we show the result of a G-VFF calculation for the strain profile in a lens-shaped InAs quantum dot embedded within GaAs. The trace of the strain is plotted in a (010) plane through the center of the dot. It shows that, to a first approximation, the InAs dot is subjected to a uniform compressive strain and the GaAs barrier is slightly expanded around the interface with the dot.

Constructing the Single-Particle Hamiltonian

We construct the single-particle Hamiltonian as

$$\hat{H} = -\frac{\beta}{2}\nabla^2 + \sum_{n\alpha}\hat{v}_\alpha(\mathbf{r} - \mathbf{R}_{n\alpha}) \quad (4)$$

where $\mathbf{R}_{n\alpha}$ is the G-VFF relaxed position of the nth atom of type α. Here $\hat{v}_\alpha(\mathbf{r})$ is a screened empirical pseudopotential for atomic type α. It contains a local part and a nonlocal, spin–orbit interaction part.

The local potential part is designed to include dependence on the local hydrostatic strain Tr(ϵ):

$$v_\alpha^{\text{loc}}(r;\epsilon) = v_\alpha^{\text{eq}}(r;0)[1 + \gamma_\alpha \text{Tr}(\epsilon)] \quad (5)$$

where the γ_α is a fitting parameter. The zero strain potential $v_\alpha^{\text{eq}}(r;0)$ is expressed in reciprocal space q as

$$v(q) = a_0(q^2 - a_1)/[a_2 e^{a_3 q^2} - 1] \quad (6)$$

where $a_{0,1,2,3}$ are fitting parameters. The local hydrostatic strain Tr(ε) for a given atom at $\mathbf{R}$ is defined as $\Omega_R/\Omega_0 - 1$, where Ω_R is the volume of the tetrahedron formed by the four atoms bonded to the atom at $\mathbf{R}$. Ω_0 is the volume of that tetrahedron in the unstrained condition. The need for explicit dependence of the atomic pseudopotential on strain in Eq. 5 results from the following: While the description in Eq. 4 of the total pseudopotential as a superposition of atomic potentials situated at specific sites, $\{R_{n\alpha}\}$, does capture the correct local symmetries in the system, the absence of a self-consistent treatment of the Schrödinger equation deprives the potential from changing in response to strain. In the absence of a strain-dependent term, the volume dependence of the energy of the bulk valence band maximum is incorrect. While self-consistent descriptions show that the volume deformation potential a_v=dE_v/d ln Ω of the valence band maximum is *negative* for GaAs, GaSb, InAs, InSb, and for all II–VI, this qualitative behavior cannot be obtained by a nonself-consistent calculation that lacks a strain-dependent pseudopotential.

In Eq. 4, the kinetic energy of the electrons has been scaled by a factor of β. The origin of this term is as follows: In an accurate description of the crystal band structure, such as the GW method, a general, spatially nonlocal potential, $V(r,r')$, is needed to describe the self-energy term. In the absence of such a term, the occupied band width of an inhomogeneous electron gas is too large compared to the exact many-body result. However, to a first approximation, the leading effects of this nonlocal potential, $V(r,r')$, can be represented by scaling the kinetic energy. This can be obtained by Fourier transforming $V(r,r')$ in reciprocal space, q, then making a Taylor expansion of q about zero. We find that the introduction of such a kinetic energy scaling, β, permits a simultaneous fit of both the effective masses and energy gaps. In this study, we fit β=1.23 for both GaAs and InAs.

The pseudopotential parameters in Eqs. 5 and 6 were fitted to the bulk band structures, experimental deformation potentials and effective masses, and first-principles calculations of the valence band offsets of GaAs and InAs. The alloy bowing parameter for the GaInAs band gap (0.6 eV) is also fitted. An example of the quality of the

pseudopotential fit for GaAs and InAs can be found in Ref. [9].

Solving the Single-Particle Hamiltonian

We have developed two techniques for solving for the eigenstates of Eq. 4. The choice of technique depends on the size of the heterostructure system being studied, i.e., the number of electrons in the system, and the desired level of accuracy. The two techniques are distinguished by: 1) the choice of the basis set for expanding the single-particle eigenstates of Eq. 4; and 2) the algorithm used to solve for the eigenstates within that particular basis. In our first approach, we expand the eigenstates in a *plane wave* basis and use the folded spectrum method to obtain eigenstates in a given energy window. In the second approach, we expand the eigenstates in a basis of *bulk Bloch orbitals* and use a Lanzcos algorithm to obtain eigenstates.

Expansion in a plane wave basis

The conventional basis set for performing calculations of the electronic band structure of periodic materials, such as bulk semiconductors, is the plane wave basis.

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}}^{E_{\text{cut}}} c_{\mathbf{G}}^i e^{i\mathbf{G}\cdot\mathbf{r}} \tag{7}$$

Within this basis, one can easily utilize fast Fourier transforms to convert the wavefunctions from a real-space to reciprocal-space representation. The matrix elements of the Hamiltonian in Eq. 4 in the basis of Eq. 7 can be written as

$$\hat{H}_{\mathbf{G},\mathbf{G}'} = \frac{1}{2}\mathbf{G}^2\delta_{\mathbf{G},\mathbf{G}'} + V_{\text{local}}(\mathbf{G}-\mathbf{G}') + V_{\text{nonlocal}}(\mathbf{G},\mathbf{G}') \tag{8}$$

The conventional variation al approach to solving for the eigenstates of Eq. 8 is to minimize the energy $\langle\psi_0|\hat{H}|\psi_0\rangle$, of the groundstate wavefuction, ψ_0, by varying its expansion coefficients, $c_{\mathbf{G},0}$. To find higher states, one needs to orthogonalize each ψ_i to all previously converged energy eigenstates below it. This orthogonalization process scales as the third power of the number of states and so only small systems with up to a few hundred electrons can be solved in this manner.

To enable us to study heterostructure systems containing 10^6 electrons, we instead "fold" the spectrum of eigenstates about a specified reference energy and hence solve[10,11] for the eigenstates of the equation

$$(\hat{H}-\epsilon_{\text{ref}})^2\psi_i = (\epsilon-\epsilon_{\text{ref}})^2\psi_i \tag{9}$$

where ϵ_{ref} is a chosen reference energy. By placing ϵ_{ref} within the band gap of the quantum dot system, and close to the valence band maximum (VBM) or conduction band minimum (CBM), one is then able to calculate the top few valence states or the bottom few conduction states, respectively. As quantum confinement effects act to lower (raise) electron (hole) levels in the quantum dot compared to the bulk, one can ensure that ϵ_{ref} falls within the band gap of the dot simply by placing it within the bulk band gap of the dot material. By applying the $\hat{H}$ operator twice, it is easy to see that the eigenstates of Eq. 9 are also eigenstates of the Hamiltonian in Eq. 4. Therefore, the process of folding the spectrum does not introduce any additional approximations compared to the conventional N^3 scaling algorithms. It simply removes the need to calculate all the low-energy eigensolutions and then orthogonalize to each of these states. By removing the need for this costly orthogonalization, the folded spectrum algorithm is able to scale linearly with the number of electrons in the system. A version of this folded spectrum code was developed[12] for parallel supercomputers, which linearly scales up to hundreds of processors and is able to handle systems containing up to 10^7 electrons. Typically, a calculation requires 20–50 plane waves per atom in the system and so these parallel calculations are effectively finding selective eigenstates of a matrix of order 10^8.

Expansion in a linear combination of Bloch bands

For larger quantum dot heterostructure systems containing several million atoms, using a plane wave basis set to expand the single-particle eigenstates becomes too computationally demanding, even when the folded spectrum method is employed. Instead, we choose to use a more physically intuitive basis set, namely a linear combination of bulk bands (LCBB):[13]

$$\psi_i(\mathbf{r}) = \sum_s \sum_{n,k} c_{s,n,k}^{(i)} u_{s,n,k}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{10}$$

where $u_{s,n,k}(\mathbf{r})$ is the cell periodic part of the bulk Bloch wavefunction for structure, s, at the nth band and the kth k-point. As these states form a physically more intuitive basis than traditional plane waves, the number of bands and k-points can be significantly reduced to keep only the physically important bands and k-points. This method was recently generalized to strained semiconductor heterostructure systems[14] and to include the spin–orbit interaction. To study an $In_xGa_{1-x}As$ quantum dot embedded within a GaAs barrier, we typically use an LCBB basis derived from a set of structures, s. This allows us to

directly include the effects of strain on the basis functions. The structures used are as follows: 1) unstrained, bulk InAs at zero pressure; 2) unstrained, bulk GaAs at zero pressure; 3) bulk InAs subjected to the strain value in the center of the InAs dot; and 4) bulk InAs subjected to the strain value at the tip of the InAs dot. For an $In_xGa_{1-x}As$ quantum dot system, where we expect the electron and hole states to be derived from around the bulk Γ point, the wavevectors, $\{k\}$, include all allowed values within a given cutoff of the zone center. For calculations of electron states, we find we only need to include the band index, n, around the Γ_{1c} point. For the hole states, we also include the three bands around the Γ_{15v} point. As the number of k-points and bands in the LCBB basis is increased, the eigenstates converge to those calculated using a converged basis, such as the above plane wave basis with a large cutoff, E_{cut}. Therefore we can use a plane wave, folded spectrum calculation as a reference to judge when our LCBB basis set is sufficient. We find that to study a typical quantum dot system, a basis set containing 10,000 bulk bands produces single-particle energies that are converged with respect to basis size, to within 1 meV.

Calculation of "Two-Body" Interactions

Using screened Hartree Fock theory, the energy associated with loading N electrons and M holes into a quantum dot can be expressed[15] as

$$\begin{aligned} E_{MN} = & \sum_i -\epsilon_{h_i} m_i + \sum_{i<j} (J_{ij}^{hh} - K_{ij}^{hh}) m_i m_j \\ & + \sum_i \epsilon_{e_i} n_i + \sum_{i<j} (J_{ij}^{ee} - K_{ij}^{ee}) n_i n_j \\ & - \sum_{ij} (J_{ij}^{eh} - K_{ij}^{eh}) n_i m_j \end{aligned} \tag{11}$$

where the electron and hole levels are denoted by $e_0,e_1,e_2,\ldots$, and $h_0,h_1,h_2,\ldots$, respectively. The n_i and m_i are the electron and hole occupation numbers, respectively, such that $\sum_i n_i = N$ and $\sum_i m_i = M$. The ϵ_i are the single-particle energies of the ith state, J_{ij} and K_{ij} are the direct and exchange Coulomb integrals between the ith and jth electronic states. For example, by using Eq. 11, in the strong confinement regime where kinetic energy effects dominate over the effects of exchange and correlation, an exciton involving an electron excited from hole state i to electron state j can be expressed as

$$E_{ij}^{exciton} = (\epsilon_{e_j} - \epsilon_{h_i}) - J_{ji}^{eh} + K_{ji}^{eh}\delta_{S,0} \tag{12}$$

The direct and exchange Coulomb energies are defined[16] as

$$\begin{aligned} J_{ijkl} &= \iint \frac{\psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_2)\psi_k^*(\mathbf{r}_1)\psi_l(\mathbf{r}_2)}{\bar{\epsilon}(\mathbf{r}_1 - \mathbf{r}_2)|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \\ K_{ijkl} &= \iint \frac{\psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_2)\psi_k^*(\mathbf{r}_2)\psi_l(\mathbf{r}_1)}{\bar{\epsilon}(\mathbf{r}_1 - \mathbf{r}_2)|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \end{aligned} \tag{13}$$

where $\bar{\epsilon}$ is connected to a phenomenological, screened dielectric function,[17] $\epsilon(\mathbf{r},\mathbf{r}',R)$, by,

$$\frac{1}{\bar{\epsilon}(\mathbf{r},\mathbf{r}'')|\mathbf{r} - \mathbf{r}''|} = \int d\mathbf{r}' \epsilon^{-1}(\mathbf{r},\mathbf{r}';R) \frac{1}{|\mathbf{r}'' - \mathbf{r}'|} \tag{14}$$

where R is the diameter of the quantum dot being studied.

RECENT APPLICATIONS

One of the long-standing problems that has held back the development of accurate models for the energy states in semiconductor quantum dot heterostructures is the need to accurately determine the size, shape, and composition of quantum dot samples. This problem is further compounded by the fact that, while it is possible to determine the shape of dots by using atomic force microscopy (AFM) *before* they have been capped with a GaAs "barrier," it is believed that the capping process itself induces the

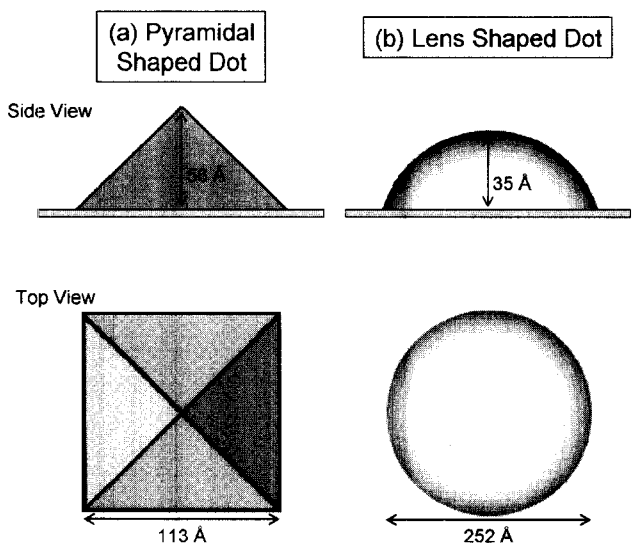


Fig. 2 The assumed geometry of the (a) pyramidal and (b) lens-shaped dots described in the sections "Pyramidal Quantum Dots: Single-Particle Electron and Hole States" and "Lens-Shaped Dots: The Effect of Changing the Shape and Composition Profile."

diffusion of gallium into the dots and diffusion of indium from the dots into the surrounding matrix. Hence AFM data for the size and shape of *uncapped* dots is only of limited use in evaluating the quality of theoretical models for *capped* dots. Therefore a major determining factor in our choice of which systems to apply the above EPM techniques to has been the continuing experimental progress in the characterization of self-assembled quantum dots. The earliest quantum dot samples were believed to contain pure InAs, pyramidal structures, with {101} facets (Fig. 2) forming 45° angles between the facets and the base. Following this interpretation of the structure, early calculations were also performed assuming a pyramidal geometry. In the section "Pyramidal Quantum Dots: Single-Particle Electron and Hole States," we compare the results of our initial calculations for pyramidal InAs quantum dots embedded within GaAs with those from other theoretical techniques. More recent characterization using cross-sectional transmission electron microscopy (TEM) and scanning tunneling microscopy (STM) measurements of *capped* dots has predicted that a more realistic dot geometry and composition is a lens-shaped In_xGa_{1-x} As quantum dot. In the section "Lens-Shaped Dots: The Effect of Changing the Shape and Composition Profile," we present results of calculations for the energy states, excitonic band gaps, and Coulomb matrix elements in lens-shaped In_xGa_{1-x} As quantum dots. These calculations were used in combination with a range of measured properties to help to determine both the geometry and composition of these quantum dots.

Pyramidal Quantum Dots: Single-Particle Electron and Hole States

The EPM techniques described in the section "Description of Pseudopotential Techniques" have been applied to the study of pyramidal self-assembled quantum dots in several publications.[17–19] The assumed pyramidal geometry is illustrated in Fig. 2a. In Fig. 3 we schematically show the single-particle energy levels that are typically calculated. The five lowest energy electron states are labeled in increasing energy from e_0 to e_4 and the three highest energy hole states are labeled in order of decreasing energy from h_0 to h_2. The single-particle wave-

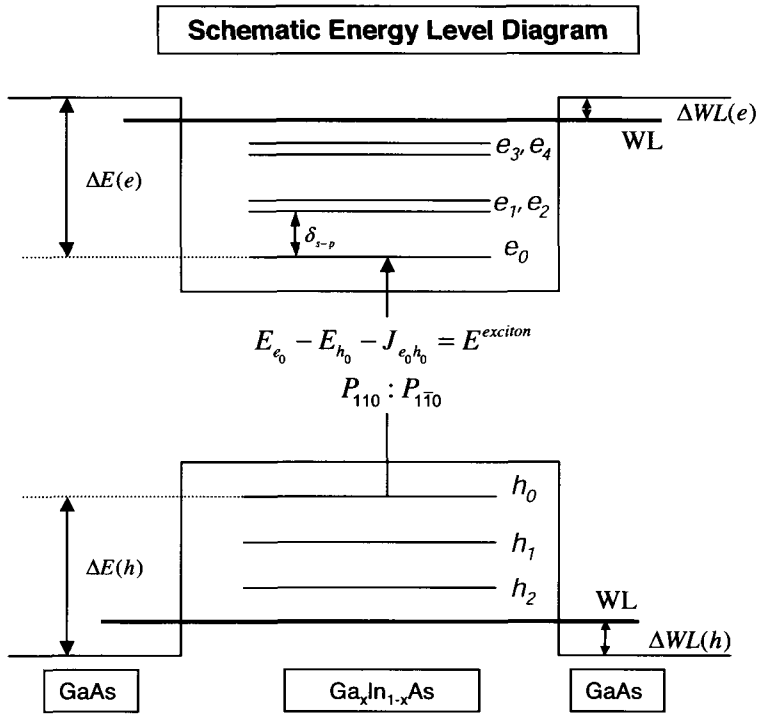


Fig. 3 A schematic representation of the single-particle energy levels in a self-assembled quantum dot. The five lowest energy electron states are labeled in increasing energy from e_0 to e_4, and the three highest energy hole states are labeled in order of decreasing energy from h_0 to h_2. The electron and hole binding energies are labeled $\Delta E(e,h)$, and the wetting layer energies with respect to the GaAs VBM and CBM are labeled $\Delta WL(h,e)$.

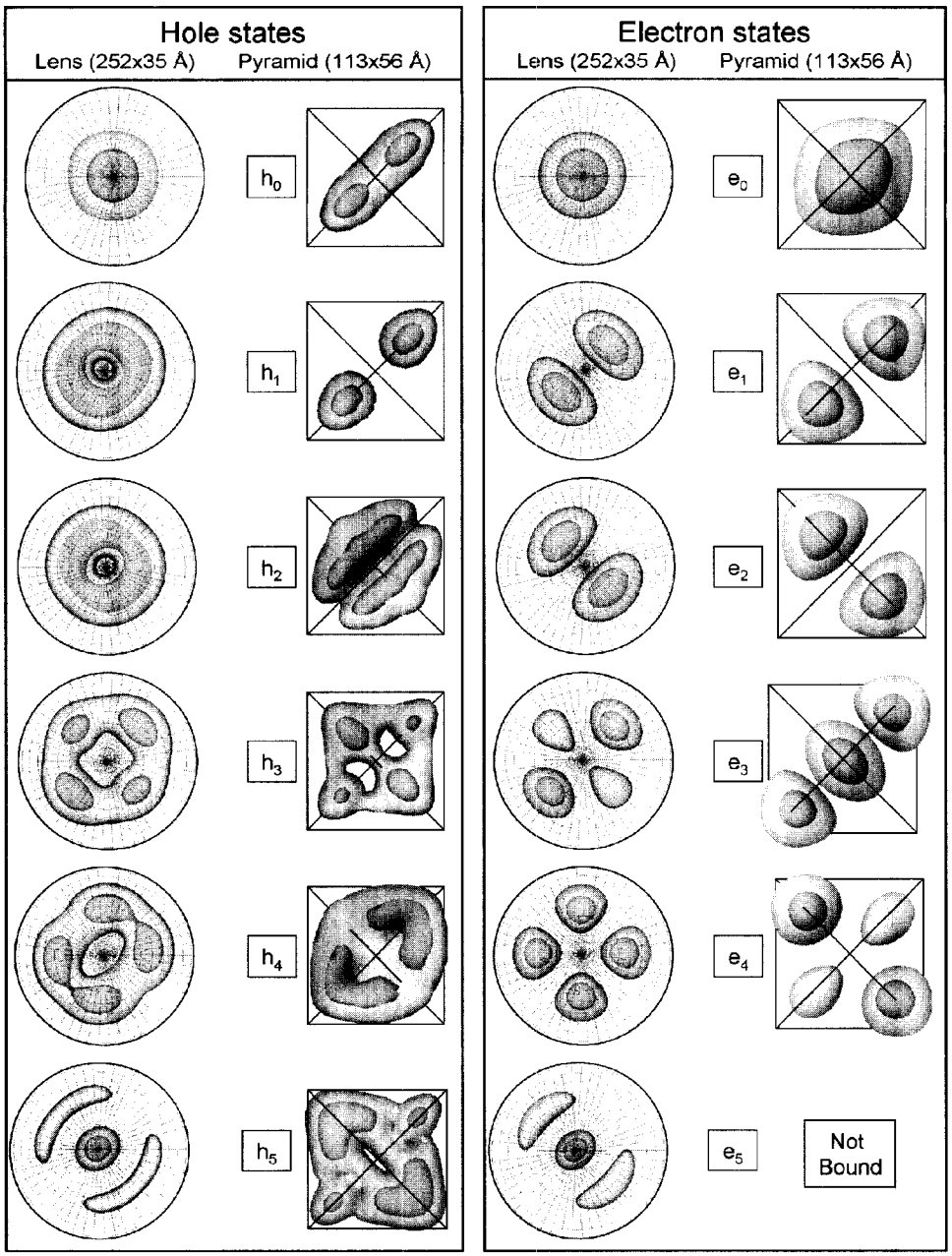


Fig. 4 Top view of the calculated electron and hole wavefunctions squared for pyramidal and lens-shaped InAs quantum dots embedded in GaAs, with bases of 252 and 113 Å and heights of 25 and 56 Å. The grey and dark grey isosurfaces represent 20% and 60% of the maximum charge density, respectively. (*View this art in color at www.dekker.com.*)

functions corresponding to each of these electron and hole states are illustrated in Fig. 4. Fig. 4 shows only the square of the envelope function, $f(\mathbf{r})$, for each of the single-particle states, i, defined as

$$f^i(\mathbf{r}) = \sum_{\mathbf{k}} c^i_{n,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{15}$$

where the $c^i_{n,\mathbf{k}}$ are the coefficients of the ith single-particle wavefunction, $\psi_i(\mathbf{r})$, when expanded in a basis of bulk Bloch orbitals representing a large number of bands, N, at the Brillouin zone center (Γ point),

$$\psi_i(\mathbf{r}) = \sum_{n=1}^{N} \sum_{\mathbf{k}} c^i_{n,\mathbf{k}} u_{n,\Gamma}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{16}$$

The expansion coefficients, $c^i_{n,\mathbf{k}}$, are obtained by projecting each of the calculated single-particle states onto this zone center basis,

$$c^i_{n,\mathbf{k}} = \langle \psi_i(\mathbf{r}) | u_{n,\Gamma}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \rangle \quad (17)$$

Plotting the envelope function squared, $f(\mathbf{r})^2$, instead of the wavefunction squared, $\psi_i(\mathbf{r})^2$, has the effect of averaging out the atomic-scale oscillations in the wavefunction, purely for graphical purposes. In all our calculations, the atomistic structure of the EPM technique automatically includes both the envelope, $f(\mathbf{r})$, and atomistic, $u_{n,\Gamma}(\mathbf{r})$, contributions to the wavefunctions. A useful intuitive guide to interpreting the single-particle electron states is to consider the eigenstates of the $\hat{L}_z$ operator.[3] The first six bound electron states corresponding to $l_z=0$, ±1, and ±2. The first state e_0, has $l_z=0$ and is commonly described as s-like as it has no nodes. The e_1 and e_2 states have $l_z=\pm1$, and are p-like with nodal planes (110) and ($\bar{1}$10). The e_3, e_4, and e_5 states have $l_z=\pm2$ and 0, respectively, and are commonly described as $d_{x^2-y^2}$, d_{xy}, and $2s$, respectively. Obviously, the square base of the pyramid reduces the symmetry of the dot from C_∞ to C_{4v}. In addition, the underlying zincblende atomistic structure further reduces the symmetry to C_{2v}. Hence, the e_0 to e_5 states correspond to the a_1, b_1, b_2, a_1, a_2, and a_1 irreducible representations of the C_{2v} group, rather than eigenstates of $\hat{L}_z$. The alignment of the e_1 and e_2p-states states along the [110] and [1$\bar{1}$0] directions results from the underlying zincblende lattice structure. Note that this simple analysis neglects the effects of the spin–orbit interaction, which further reduces the symmetry from the C_{2v} group to a double group with the same single representation for all the states. In our calculations, the spin—orbit interaction is included, but it produces no significant effects for the electron states. For the pyramidal dot studied in Fig. 4, the small size of the dot (113 Å base) and the relatively light effective mass of electrons in InAs (0.023 m_0) results in a large quantum confinement of the electron states which, in turn, pushes these states up in energy so that only the five states e_0 to e_4 are bound in the pyramidal dot. Here our definition of a bound electron (hole) state is that its energy is lower (higher) than the bulk GaAs CBM (VBM).

Fig. 4 also shows calculated envelope functions squared for the hole states in the pyramidal InAs/GaAs quantum dot. As there is a strong mixing between the original bulk Bloch states with Γ_{8v} and Γ_{7v} symmetry, the hole states cannot be interpreted as the solutions of a single-band Hamiltonian. The larger effective mass for holes results in a reduced quantum confinement of the hole states and consequently, many more bound hole states. Only the six bound hole states with the highest energy are shown in Fig. 4.

In Fig. 5, we show how the energy of the lowest 4 electron states and highest 4 hole states in a pyramidal InAs quantum dot depends on the size of the dot. It shows that, as the size of the dot increases, the quantum confinement of the electrons and holes in the dot is reduced and therefore the electron levels decrease in energy and the hole levels increase in energy. Reducing the quantum confinement by increasing the size of the dots also acts to increase the number of bound states in the dot. For example, the e_3 state is bound for dots with a base size greater than 90 Å, but is higher than the wetting layer energy for dots with a base less than 90 Å, and is therefore effectively unbound. It is interesting to compare these results with earlier calculations by: 1) Grundman et al.[20] and Cusak et al.[21] who applied single band effective techniques to pyramidal InAs dots with base sizes ranging from 60 to 160 Å and found only a single bound state; and 2) eight-band $\mathbf{k}\cdot\mathbf{p}$ calculations by Jiang and Singh[22] and Pryor,[23] who found three bound electron states over a similar range of sizes. Therefore we conclude that obtaining even qualitative estimates for the quantum confinement energies in these systems requires a multiband technique of at least eight bands.

Lens-Shaped Dots: The Effect of Changing the Shape and Composition Profile

In the section "Pyramidal Quantum Dots: Single-Particle Electron and Hole States," we discussed the single-particle electron and hole states in an idealized, pure InAs, pyramidal quantum dot. Recently, considerable evidence has emerged that, in fact, a more realistic geometry for self-assembled InAs/GaAs quantum dots is that of a lens-shaped dot[24–31] [see Fig. 2b]. In addition to modifying the predicted shape of self-assembled quantum dots, recent measurements[32–35] also indicate that the composition of these dots differs from the pure InAs that was originally assumed. In Table 1, we show the results of calculations for a pure InAs, lens-shaped quantum dot, with a base of 252 Å and a height of 35 Å, embedded within GaAs [column (a)]. Table 1 also shows the experimentally measured splittings of the electron levels, the electron–electron and electron–hole Coulomb energies, the magnetic field dependence, and the excitonic band gap measured in Refs. [26] and [30]. The agreement between the measured energy level spacings, Coulomb energies and magnetic field response with our theoretical lens-shaped model is generally good. Both the model and experiment find: 1) a large spacing, δ_{sp} ($\sim$50–60 meV), between the s-like e_0 state and the p-like e_1 state; 2) a small spacing, δ_{pp} ($\sim$3 meV), between the two p-like e_1 and e_2 states; and 3) a large spacing ($\sim$55 meV) between the p-like e_2 state and the d-like e_3 state.

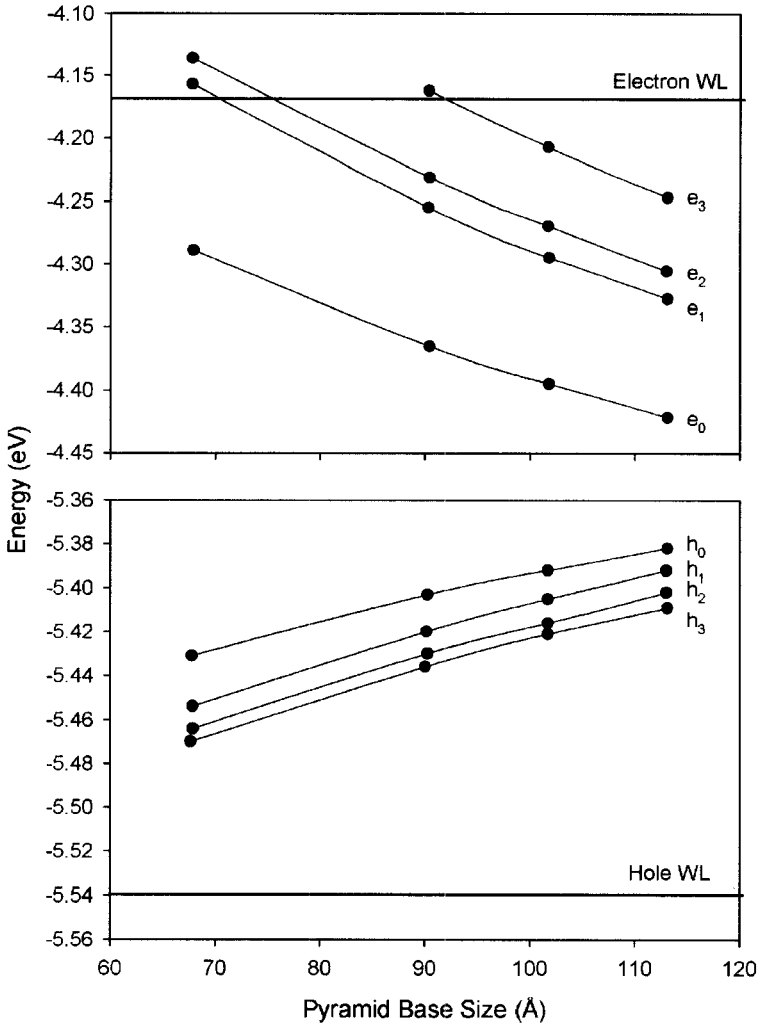


Fig. 5 The dependence of the energy of the four lowest electron states and four highest hole states on the base size of a pyramidal InAs quantum dot. The horizontal lines show the energy of the groundstate electron and hole energies in a 1-ML-thick InAs WL.

These electron level spacings are similar to those found for pyramidal quantum dots[18] [see Table 1, column (g)]. However, because of the lower pyramidal symmetry, the spacings of the two p-like and d-like states, δ_{pp} and δ_{dd}, are larger (26 and 23 meV, respectively). Both the model and experiment also find similar values for the Coulomb energies, $J(e_0e_0)$ and $J(e_0h_0)$ ($\sim$25 meV).

The calculated hole binding energy of $\Delta E(h)=193$ meV is in good agreement with those of Berryman et al.[36] ($\sim$240 meV) and Itskevich et al.[37] ($\sim$250 meV). The calculated electron–electron and electron–hole Coulomb energies are in reasonable agreement with those extracted from Refs. [26] and [30]. For the integrals $J^{ee}_{e_0e_0}$, $J^{ee}_{e_0e_1}$, $J^{ee}_{e_1e_1}$, and $J^{eh}_{e_0h_0}$, we calculate values of 31, 25, 25, and 37, respectively, compared to measured values of 23, 24, 18, and 33.3 meV. The calculated ratios of absorption intensities for light polarized along [110] and [1$\bar{1}$0] directions, defined as

$$\lambda = \frac{P_{[110]}}{P_{[1\bar{1}0]}} = \frac{\langle\psi_{e_0}|r_{[110]}|\psi_{h_0}\rangle^2}{\langle\psi_{e_0}|r_{[1\bar{1}0]}|\psi_{h_0}\rangle^2} \tag{18}$$

are $\lambda=1.03$ and 1.2, respectively, for the e_0-h_0 recombination in lens and pyramidal shaped, pure InAs dots.

In the lens-shaped dot, we find a difference in the average positions of the h_0 and e_0 states, d_{h_i,e_j}, of around 1 Å. This is smaller than the value that we calculated for a pyramidal quantum dot, where we find the hole approximately 3.1 Å higher than the electron.

In summary, the assumed lens-shaped geometry with a pure InAs composition produces a good agreement

Table 1 Calculated single particle electron and hole energy level spacings, electron and hole binding energies, $\Delta E(e,h)$, electron–electron and electron–hole Coulomb energies, excitonic band gap (all in meV), exciton dipole moment and polarization anisotropy for lens-shaped and pyramidal $Ga_xIn_{1-x}As$ quantum dots embedded within GaAs

	Lens calculations						Pyramid calculation	
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	
Geometry (Å)	**252 × 35**	**275 × 35**	**252 × 25**	**252 × 35**	**252 × 35**	**275 × 35**	**200 × 100**	
% Ga at base, tip	**0, 0**	**0, 0**	**0, 0**	**15, 15**	**30, 0**	**15, 15**	**0, 0**	**Lens experiment[26,30]**
e_1-e_0	65	57	69	58	64	52	108	50
e_3-e_2	68	61	67	60	63	57	64	48
e_2-e_1	2	2	2	2	3	2	26	2
e_2-e_1 (15 T)	20	20	18	21	20	17		19
e_4-e_3	4	3	4	4	3	1	23	
h_0-h_1	8	12	16	13	14	11	15	
h_1-h_2	7	6	5	5	6	5	20	
h_2-h_3	6	10	14	13	14	9	1	
$\Delta E(e)$	271	258	251	209	192	204	171	
$\Delta E(h)$	193	186	174	199	203	201	198	
$J_{e_0e_0}$	31	29	32	29	31	28	40	23
$J_{e_0e_1}$	25	24	26	24	24	24	35	24
$J_{e_1e_1}$	25	24	26	25	24	26	36	~18
$J_{h_0h_0}$	30	27	39	32	28	30	31	
$J_{e_0h_0}$	30	28	35	31	29	29	31	33.3
E_{gap}	1032	1016	1131	1080	1125	1083	1127	1098
d_{e_0,h_0} (Å)	0.16	−0.37	0.5	0.5	1.2	0.5	3.1	
$\lambda = P_{110} : P_{1\bar{1}0}$	1.03	1.01	1.04	1.05	1.08	1.08	1.20	

with measured level splitting, Coulomb energies, and magnetic field dependence. Detailed inspection of the remaining differences reveals that the calculations systematically *overestimate* the splittings between the single-particle electron levels (δ_{sp}: 65 vs. 50 meV, δ_{pd}:68 vs. 48 meV) and *underestimate* the excitonic band gap (1032 vs. 1098 meV).

Pure InAs dots: The effects of lens shape and size

Focusing only on the lens shape, we examine the effect of changing the height and base of the assumed geometry. Calculations were performed on similar lens-shaped, pure InAs dots where 1) the base of the dot was increased from 252 to 275 Å, while keeping the height fixed at 35 Å, [column (b)] and 2) the height of the dot was decreased from 35 to 25 Å, while keeping the base fixed at 252 Å, [column (c)]. These show that decreasing the height of the dot increases the quantum confinement and hence increases the splittings of the electron and hole levels (δ_{sp}: from 65 to 69 meV and δ_{h_0,h_1}: from 8 to 16 meV). Decreasing the height of the dot also acts to increase the excitonic band gap from 1032 to 1131 meV by pushing up the energy of the electron levels and pushing down the hole levels. Conversely, increasing the base of the dot decreases both the splittings of the single-particle levels (δ_{sp}: from 66 to 61 meV) and the band gap (1032–1016 eV). These small changes in the geometry of the lens-shaped dot have only a small effect on electronic properties that depend on the shape of the wavefunctions. The electron–electron and electron–hole Coulomb energies remain relatively unchanged, the magnetic field induced splitting remain at 20 meV, the polarization anisotropy (λ) remains close to 1.0, and the excitonic dipole, d_{h_i,e_j}, remains negligible. In summary, reducing either the height or the base of the dot increases quantum confinement effects and hence increases energy spacings and band gaps, while not significantly effecting the shape of the wavefunctions.

Interdiffused In(Ga)As/GaAs lens-shaped dots

We next investigate the effect of changing the composition of the quantum dots, while keeping the geometry fixed. Recently, there have been several experiments[33,34,38]

suggesting that a significant amount of Ga diffuses into the nominally pure InAs quantum dots during the growth process. We investigate two possible mechanisms for this Ga in-diffusion: 1) Ga diffuses into the dots during the growth process from all directions producing a dot with a uniform Ga composition $Ga_xIn_{1-x}As$; and 2) Ga diffuses up from the substrate, as suggested in Ref. [38]. To investigate the effects of these two methods of Ga in-diffusion on the electronic structure of the dots, we compare pure InAs dots embedded in GaAs with $Ga_xIn_{1-x}As$, random alloy dots embedded in GaAs, where the Ga composition, x, 1) is fixed at 0.15, [column (d)] and 2) varies linearly from 0.3 at the base to 0 at the top of the dot, [column (e)].

Table 1 shows that increasing the amount of Ga in the dots acts to decrease the electron level spacings (δ_{sp}: from 65 to 58 for $x=0.15$). It also acts to increase the excitonic band gap from 1032 to 1080 and 1125 meV, respectively. The electron binding energy, $\Delta E(e)$, is decreased by the in diffusion of Ga (from 271 to 209 and 192 meV), while the hole binding energy, $\Delta E(h)$, is relatively unaffected. This significant decrease in the electron binding energy considerably improves the agreement with experiments on other dot geometries.[36,39]

As with changing the size of the dots, we find that Ga in-diffusion has only a small effect on properties that depend on the shape of the wavefunctions. The calculated electron–electron and electron–hole Coulomb energies are almost unchanged, while the average separation of the electron and hole, d_{h_i,e_j}, increases from 0.16 to 0.5 and 1.2 Å and the polarization ratio, λ, and magnetic field response are also unchanged.

Table 1 shows that the dominant contribution to the increase in the excitonic band gap and reduction in electron binding energy, results mostly from an increase in the energy of the *electron* levels as the Ga composition is increased. This can be understood by considering the electronic properties of the bulk $Ga_xIn_{1-x}As$ random alloy. The unstrained valence band offset between GaAs and InAs is $\sim$50 meV,[40] while the conduction band offset in $\sim$1100 meV and hence changing the Ga composition, x, has a large effect on the energy of the electron states and only a small effect on the hole states.

In summary, the effect of Ga in-diffusion is to reduce the spacing of the electron levels while significantly increasing their energy and hence increasing the band gap. We find that only the average Ga composition in the dots is important to their electronic properties. Whether this Ga is uniformly or linearly distributed throughout the dots has a negligible effect.

The effects of changing the *geometry* of the lens-shaped, pure InAs dots on the single-particle energy levels can be qualitatively understood from single-band, effective mass arguments. These predict that decreasing any dimension of the dot increases the quantum confinement and hence, the energy level spacings and the single-particle band gap will increase. Note that as the dominant quantum confinement in these systems arises from the vertical confinement of the electron and hole wavefunctions, changing the height has a stronger effect on the energy levels than changing the base. In this case, decreasing the height by 10 Å has a much stronger effect on the energy spacings and on the band gap than increasing the base by 23 Å.

As increasing (decreasing) the dimensions of the dot acts to decrease (increase) both the level spacings and the gap, it is clear that changing the dot geometry alone will not significantly improve the agreement with experiment as this requires a simultaneous *decrease* in the energy level splittings and *increase* in the band gap. However, Ga in-diffusion into the dots acts to *increase* the band gap of the dot while decreasing the energy level spacings. Table 1 shows that adopting a geometry with a base of 275 Å and a height of 35 Å and a uniform Ga composition of $Ga_{0.15}In_{0.85}As$ produces the best fit to the measurements in Refs. [26] and [30].

CONCLUSION

In conclusion, we believe we have successfully demonstrated that modern empirical pseudopotential techniques, coupled with atomistic strain calculations, provide a powerful technique for calculating the electronic structure of semiconductor nanostructures containing several million atoms.

Our results strongly suggest that to obtain accurate agreement between theoretical models and experimental measurements for standard lens-shaped quantum dots, one needs to adopt a model of the quantum dot that includes some Ga in-diffusion within the quantum dot. When 15% Ga in-diffusion is included, we obtain an excellent agreement between state-of-the-art multiband pseudopotential calculations and experiments for a wide range of electronic properties. We are able to fit/predict most observable properties to an accuracy of ± 5 meV, which is sufficient to make predictions of both the geometry and composition of the dot samples.

The techniques described above have recently been extending to describing the decay of multiple excitons in semiconductor quantum dots. For more details, see Refs. [9] and [41].

ACKNOWLEDGMENTS

The author would like to thank Alex Zunger, Lin-Wang Wang, Andrew Canning, Alberto Franceschetti, and John Shumway for collaborations on the above projects.

The work performed by the author at the National Renewable Energy Laboratory was supported by the Department of Energy Basic Energy Sciences, Division of Materials Science under contract No. DE-AC36-98-GO10337. The work performed by the author at Lawrence Livermore National Laboratory was performed under the auspices of the DOE by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

REFERENCES

1. Yoffe, A. Low-dimensional systems: Quantum size effects and electronic properties of semiconductor microcrystallites (zero-dimensional systems) and some quasi-two-dimensional systems. Adv. Phys. **2002**, *51*, 799.
2. Bimberg, D.; Grundmann, M.; Ledentsov, N. *Quantum Dot Heterostructures*; Wiley: Chichester, 1998.
3. Jacak, L.; Hawrylak, P.; Wojs, A. *Quantum Dots*; Springer: Berlin, 1998.
4. Wang, L.; Williamson, A.; Zunger, A.; Jiang, H.; Singh, J. Comparison of the kp and direct diagonalization approaches to the electronic structure of InAs/GaAs quantum dots. Appl. Phys. Lett. **2000**, *76*, 339.
5. Keating, P. Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure. Phys. Rev. **1966**, *145*, 637.
6. Pryor, C.; Kim, J.; Wang, L.; Williamson, A.; Zunger, A. Comparison of two methods for describing the strain profiles in quantum dots. J. Appl. Phys. **1998**, *83*, 2548.
7. Martin, R. Elastic properties of zns structure semiconductors. Phys. Rev., B **1970**, *1*, 4005.
8. Williamson, A. *Energy States in Quantum Dots*; World Scientific: Singapore, 2002. Chapter 1.
9. Williamson, A.; Grossman, J.; Hood, R.; Puzder, A.; Galli, G. Quantum Monte Carlo calculations of nanostructure optical properties: Application to silicon quantum dots. Rev. Rev. Lett. **2002**, *89*, 196803.
10. Wang, L.; Zunger, A. *Semiconductor Nanoclusters*; Elsevier Science: Amsterdam, 1996.
11. Wang, L.W.; Zunger, A. Solving Schrodingers equation around a desired energy-application to silicon quantum dots. J. Phys. Chem. **1994**, *100*, 2394.
12. Canning, A.; Wang, L.W.; Williamson, A.; Zunger, A. Parallel empirical pseudopotential electronic structure calculations for million atom systems. J. Comp. Physiol. **2000**, *160*, 29.
13. Wang, L.; Franceschetti, A.; Zunger, A. Million-atom pseudopotential calculation of gamma–x mixing in GaAs/AlAs superlattices and quantum dots. Phys. Rev. Lett. **1997**, *78*, 2819.
14. Wang, L.-W.; Zunger, A. Linear combination of bulk bands method for large-scale electronic structure calculations on strained nanostructures. Phys. Rev., B **1999**, *59*, 15806.
15. Franceschetti, A.; Williamson, A.; Zunger, A. Addition spectra of quantum dots: The role of dielectric mismatch. J. Phys. Chem. **2000**, *104*, 3398.
16. Franceschetti, A.; Zunger, A. Direct pseudopotential calculation of exciton coulomb and exchange energies in semiconductor quantum dots. Phys. Rev. Lett. **1997**, *78*, 915.
17. Williamson, A.; Zunger, A. Effect of interfacial states on the binding energies of electrons and holes in InAs/GaAs quantum dots. Phys. Rev., B **1998**, *58*, 6724.
18. Kim, J.; Wang, L.; Zunger, A. Comparison of the electronic structure of InAs/GaAs pyramidal quantum dots with different facet orientations. Phys. Rev., B **1998**, *57*, R9408.
19. Wang, L.; Kim, J.; Zunger, A. Electronic structures of [110]-faceted self-assembled pyramidal InAs/GaAs quantum dots. Phys. Rev., B **1999**, *59*, 5678.
20. Grundmann, M.; Stier, O.; Bimberg, D. InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons, electronic structure. Phys. Rev., B **1995**, *52*, 11969.
21. Cusak, M.; Briddon, P.; Jaros, M. Electronic structure of InAs/GaAs self-assembled quantum dots. Phys. Rev., B **1996**, *54*, 2300.
22. Jiang, H.; Singh, J. Conduction band spectra in self-assembled InAs/GaAs dots: A comparison of effective mass and an eight-band approach. Appl. Phys. Lett. **1997**, *71*, 3239.
23. Pryor, C. Eight-band calculations of strained InAs/GaAs quantum dots compared with one-, four-, six-band approximations. Phys. Rev., B **1998**, *57*, 7190.
24. Drexler, H.; Leonard, D.; Hansen, W.; Kotthaus, J.; Petroff, P. Spectroscopy of quantum levels in charge-tunable InGaAs quantum dots. Phys. Rev. Lett. **1994**, *73*, 2252.
25. Medeiros-Ribeiro, G.; Leonard, D.; Petroff, P. Electron and hole energy levels in InAs self-assembled quantum dots. Appl. Phys. Lett. **1995**, *66*, 1767.
26. Fricke, M.; Lorke, A.; Kotthaus, J.; Medeiros-Ribeiro, G.; Petroff, P. Shell structure and elec-

tron–electron interaction in self-assembled InAs quantum dots. Europhys. Lett. **1996**, *36*, 197.
27. Miller, B. Few-electron ground states of charge-tunable self-assembled quantum dots. Phys. Rev., B **1997**, *56*, 6764.
28. Warburton, R.; Durr, C.; Karrail, K.; Kotthaus, J.; Medeiros-Ribeiro, G.; Petroff, P. Charged excitons in self-assembled semiconductor quantum dots. Phys. Rev. Lett. **1997**, *79*, 5282.
29. Schmidt, K.; Medeiros-Robeiro, G.; Garcia, J.; Petroff, P. Size quantization effects in InAs self-assembled quantum dots. Appl. Phys. Lett. **1997**, *70*, 1727.
30. Warburton, R.; Miller, B.; Durr, C.; Bodefeld, C.; Karrai, K.; Kotthaus, J.; Medeiros-Ribeiro, G.; Petroff, P.; Huant, S. Coulomb interactions in small charge-tunable quantum dots: A simple model. Phys. Rev., B **1998**, *58*, 16221.
31. Schmidt, K.; Medeiros-Robeiro, G.; Petroff, P. Photoluminescence of charged InAs self-assembled quantum dots. Phys. Rev., B **1998**, *58*, 3597.
32. Yang, W.; Lee, H.; Johnson, J.; Sercel, P.; Norman, A. Electronic structure of self-organized InAs/GaAs quantum dots bounded by (136) facets. Phys. Rev., B **2000**, *64*, 2784.
33. Metzger, T.; Kegel, I.; Paniago, R.; Peisl, J. Grazing incidence x-ray scattering: An ideal tool to study the structure of quantum dots. J. Phys., D, Appl. Phys. **1999**, *32*, A202.
34. Garcia, J.; Medeiros-Ribeiro, G.; Schmidt, K.; Ngo, T.; Feng, J.; Lorke, A.; Kotthaus, J.; Petroff, P. Intermixing and shape changes during the formation of InAs self-assembled quantum dots. Appl. Phys. Lett. **1997**, *71*, 2014.
35. Rubin, M.; Medeiros-Ribeiro, G.; O'Shea, J.; Chin, M.; Lee, E.; Petroff, P.; Narayanamurti, V. Imaging and spectroscopy of single InAs self-assembled quantum dots using ballistic electron emission microscopy. Phys. Rev. Lett. **1996**, *77*, 5268.
36. Berryman, K.; Lyon, S.; Segev, M. Electronic structure and optical behavior of self-assembled InAs quantum dots. J. Vac. Sci. Technol., B **1997**, *15* (4), 1045.
37. Itskevich, I.; Skolnick, M.; Mowbray, D.; Trojan, I.; Lyapin, S.; Wilson, L.; Steer, M.; Hopkinson, M.; Eaves, L.; Main, P. Excited states and selection rules in self-assembled InAs/GaAs quantum dots. Phys. Rev., B **1999**, *60*, R2185.
38. Fry, P.; Itskevich, I.; Mowbray, D.; Skolnick, M.; Finley, J.; Barker, J.; O'Reilly, E.; Wilson, L.; Larkin, I.; Maksym, P.; Hopkinson, M.; Al-Khafaji, M.; David, J.; Cullis, A.; Hill, G.; Clark, J. Inverted electron–hole alignment in InAs–GaAs self-assembled quantum dots. Phys. Rev. Lett. **2000**, *84*, 733.
39. Tang, Y.; Rich, D.; Mukhametzhanov, I.; Chen, P.; Hadhukar, A. Self-assembled InAs/GaAs quantum dots studied with excitation dependent cathodoluminescence. J. Appl. Phys. **1998**, *84*, 3342.
40. Wei, S.; Zunger, A. Calculated natural band offsets of all II–VI and III–V semiconductors: Chemical trends and the role of cation d orbitals. Appl. Phys. Lett. **1998**, *72*, 2011.
41. Williamson, A.; Franceschetti, A.; Zunger, A. Multi-excitons in self-assembled InAs/GaAs quantum dots: A pseudopotential, many-body approach. Europhys. Lett. **2001**, *53*, 59.

Quantum Dots, Self-Formed: Structural and Optical Characterization

Shun-ichi Gonda
Fukui University of Technology, Fukui, Japan

Hajime Asahi
Osaka University, Osaka, Japan

INTRODUCTION

There are various methods for the fabrication of quantum dots (QD).[1,2] Among them, the self-formation method is one of the most useful methods, because it causes little damage, creates dots of high density, and the fabrication process is rather simple. In this article, two fabrication methods, i.e., the S–K mode growth method and the composition modulation of short-period superlattices, are introduced. For samples fabricated by these methods, structural characterization with atomic force microscopy (AFM), scanning tunneling microscopy/spectroscopy (STM/STS), magnetic force microscopy (MFM), and transmission electron microscopy (TEM) are described. Furthermore, as optical characterization methods, photoluminescence (PL), time-resolved photoluminescence (TRPL), and electroluminescence (EL) are mentioned and the several results are shown.

FABRICATION METHODS

Prior to the description of the characterization,[3,4] we briefly describe the fabrication methods for self-formed QDs. A frequently used method is the so-called "Stranski–Krastanov (S–K) mode growth method." In this method, a small amount of materials having a lattice constant different from that of the substrate are supplied at appropriate temperatures on a substrate (e.g., GaAs). Because of the difference of surface energy and the strain, projecting dots of supplied materials are formed on the substrate. A characteristic feature of this method is dot formation on the flat surfaces, shown schematically in Fig. 1a. The typical density of dots made by the S–K mode growth is around 10^9–10^{10} cm^{-2}. Examples of materials combination are InAs on GaAs, InGaAs on GaAs,[5–8] and InGaN on GaN.[9]

As an example, our growth process of InAs dots on GaAs using metal organic molecular beam epitaxy (MOMBE) is as follows:[10] After the growth of GaAs buffer layer, triethylgallium (TEGa) and trisdimethlaminoarsenic (TDMAAS) are supplied at an appropriate substrate temperature. The formation of InAs dots can be monitored by observing the reflection high-energy electron diffraction (RHEED) pattern changes from streaky to spotty. Other growth processes are shown in Refs. [5–9].

Another method is the "composition modulation of short-period superlattices." In this method,[11] e.g., short-period GaP/InP superlattice layers are grown on GaAs substrate. The state of grown layers depends on the orientation of GaAs substrates. With this method, wire structures are grown on (100) GaAs substrates[11] and CuPt-type GaP/InP superlattices are grown on (111) GaAs.[12] On (*N*11) GaAs substrate, dot structures are self-formed.[12] In this case, dots are formed in the grown layer. Therefore the surface is flat, as shown in Fig. 1b. The density of dots is around 10^{11} cm^{-2}. Examples of material combination are InP/GaP on GaAs and GaAs/InAs on InP.

Here the fabrication procedure of dots by the above method is described in more detail. Two types of superlattices are grown on Si-doped GaAs(311)A substrates by gas source molecular beam epitaxy. One is $(GaP)_n(InP)_m$, where n and m represent the number of the monolayer of each compound. (GaP)(InP) means that the GaP layer was grown first on the substrate and the InP layer was grown next. The other type of sample $(InP)_n(GaP)_m$ was formed in the reverse order. Elemental Ga, In, Si, and thermally cracked arsine (AsH_3) and phosphine (PH_3) were used as sources.

STRUCTURAL CHARACTERIZATION

Structure of self-formed quantum dots can be characterized by various methods such as atomic force microscopy (AFM), scanning tunneling microscopy (STM), electron microscopy, and electron diffraction.

AFM, STM, and STS

AFM utilizes the atomic force between a sample surface and a probe.[13] Therefore the obtained image is almost

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013810

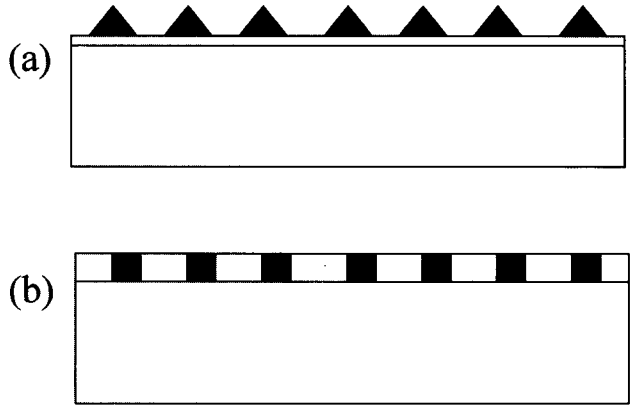


Fig. 1 Schematic view of (a) QDs fabricated by the S–K mode growth method and (b) QDs fabricated by the composition modulation method.

independent of the electric properties of materials. In STM, the height position of the probe is controlled to keep the tunnel current between sample and probe constant, and the variation of probe is displayed.[14] Hence the obtained image reflects the electrical properties of materials. In scanning tunneling spectroscopy (STS), the position of the probe is fixed and the current–voltage characteristics are measured. These characteristics reflect the electronic structures of materials.

The shape and position of QDs fabricated by the S–K mode growth method are well characterized by AFM. Fig. 2 shows AFM image for (a) InAs dots as grown and for (b) the surface where smaller dots are eliminated by etching.[15] The diameter, height, density, and distribution of dots can be measured. The dependence of dot structures on the growth conditions is investigated via AFM.[16,17] On the other hand, in the case of QDs made by composition modulation, AFM shows that the surface is flat, and information on the size and distribution of dots are not obtained.

In some special cases, such as InAs dots including Mn atoms, magnetic force microscopy (MFM) can be used.[18] For MFM, batch-microfabricated silicon probes were magnetically sensitized by sputter coating with ferromagnetic materials. The tip is scanned tens or hundreds of nanometers above the sample, thus avoiding contact. Magnetic field gradients exert a force on the tip's magnetic moment, and monitoring the tip/cantilever response gives a magnetic force image. Fig. 3 shows AFM and MFM images of the InAs dots including Mn atoms.[19]

In the case of QDs by composition modulation, STM is a powerful tool for characterization. The STM image of the sample $(GaP)_{1.5}(InP)_{1.88}$ on GaAs(311)A substrate is shown in Fig. 4a.[20] In the STM image reflecting the electrical properties, shape and distribution can be observed. The self-formed dots are aligned along two perpendicular directions, [0–11] and [233], and distributed side by side. The lateral period (size) of dots (bottom to bottom in the STM image) is about 20 nm (~20 nm along the [0–11] direction, ~19 nm along the [233] direction). The distribution (size fluctuation) is about ±10%. The dot density is of the order 10^{11} cm^{-2}. The bright and dark areas correspond to the InP-rich and GaP-rich regions, respectively, based on STS measurement.

Fig. 4b shows a dI/dV vs. bias voltage curve at points A, B, and C in the STS measurement. The voltage width for $dI/dV=0$ corresponds to the potential change, i.e., the

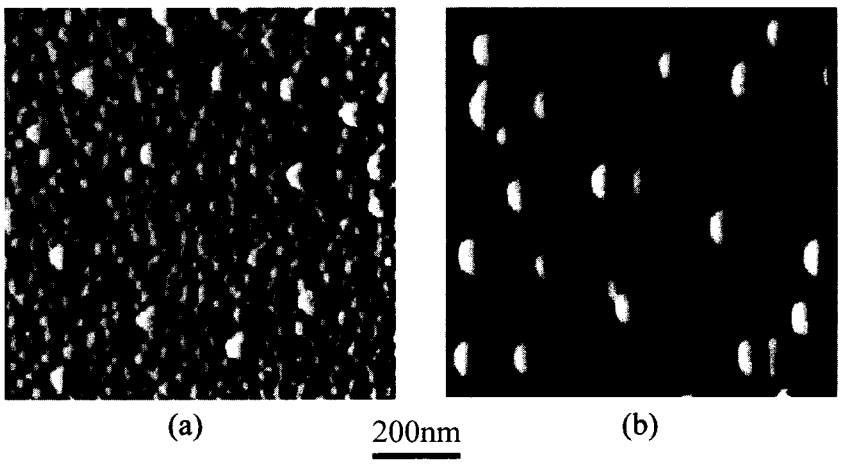


Fig. 2 AFM images for (a) the InAs dots as grown by S–K mode growth and for (b) the surface where smaller dots are eliminated by etching. (From Ref. [15].) (*View this art in color at www.dekker.com.*)

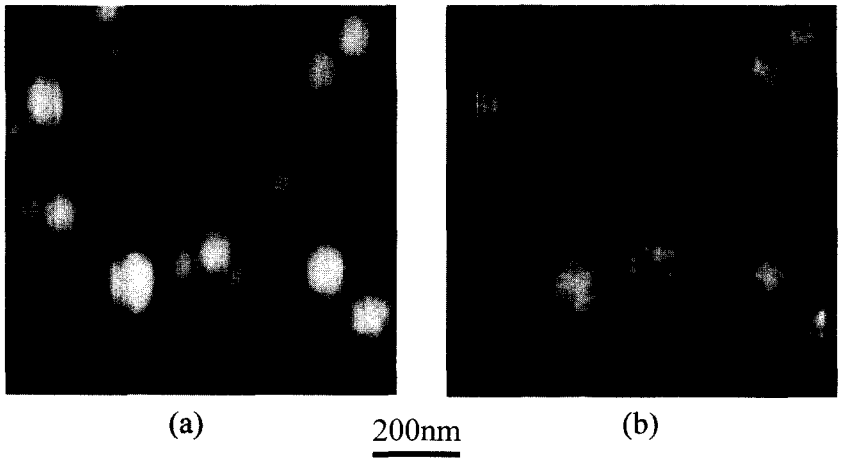


Fig. 3 (a) AFM and (b) MFM images for the InAs dots including Mn atoms. (From Ref. [19].) (*View this art in color at www.dekker.com.*)

band gap change of the materials. The periodic structure of the voltage width shown in Fig. 4c is not attributable to the geometrical height change, but to the band gap change, because the surface is almost flat as described before.

In the case of the $(InP)_{1.88}(GaP)_{1.5}$ sample, nearly the same dot structures are observed in the STM measurement. The difference is that the lateral period is about 22 mm (~22 nm along the [0–11] direction, ~20 nm along the [233] direction). This is 10% larger than that of the $(GaP)_{1.5}(InP)_{1.88}$ sample. The size distribution is also about ±10%, which is nearly the same as $(GaP)_{1.5}(InP)_{1.88}$ sample. The difference in the lateral period (size) between these two types of QDs was confirmed for several samples.

STM can be also utilized for the characterization of QDs made by the S–K mode growth. (Please see Refs. [21,22].)

Electron Microscopy

Inner structures of dots can be observed by using cross-sectional transmission electron microscope (TEM). It is important to carefully prepare specimens without causing damage. First, the sample is sliced or cleaved into bars. The bars, the backside of which is supported with Ti holder, are mounted on glass or metal with carbon epoxy so as to make the cross section upside. The bars are thinned by polishing and Ar milling to less than 10 nm. Fig. 5a shows a cross-sectional TEM lattice image of the $(GaP)_{1.5}(InP)_{1.88}$ sample grown on (311)A GaAs substrate by composition modulation method.[3] This shows the lateral periodic thickness undulation, as shown schematically in Fig. 5b, although GaP/InP superlattice structures are essentially preserved.

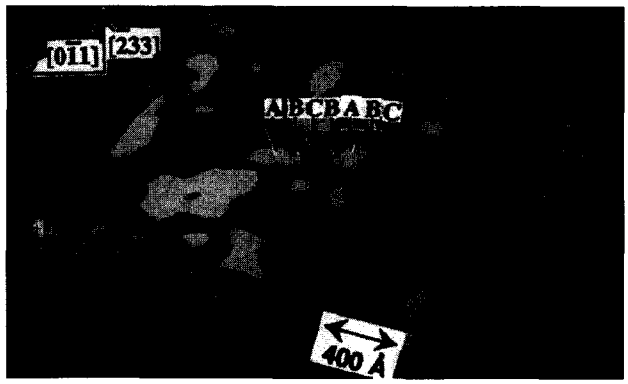
Characterizations of QDs made by the S–K mode growth with cross-sectional TEM images are reported.[23,24]

Electron Diffraction

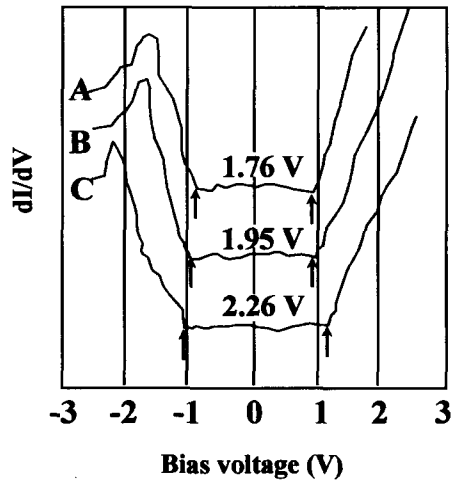
TEM diffraction is also used for the structural characterization. Fig. 6 shows TEM diffraction patterns for the $(GaP)_1(InP)_1$ superlattices grown on (a) GaAs(100) and (b) GaAs(111)B. Incident electron beam is parallel to the [111] direction in the case of Fig. 6a and to the [211] direction in Fig. 6b.[25] The sharp spots (0 1/2 1/2) and (0 −1/2 −1/2) were clearly observed in Fig. 6a. This indicates that the superstructure with a period of twice the usual lattice constant is formed in the [0 1 1] direction. On the other hand, the sharp superstructure spots (1/2 1/2 1/2) were observed in Fig. 6b. This indicates that the CuPt-type long-range ordering, i.e., $(GaP)_1(InP)_1$ superlattices, is formed in the [111] direction. These observations show that the growth on the (100) and (111) substrates does not form quantum dots by the composition modulation method.

X-ray Scattering and Diffraction

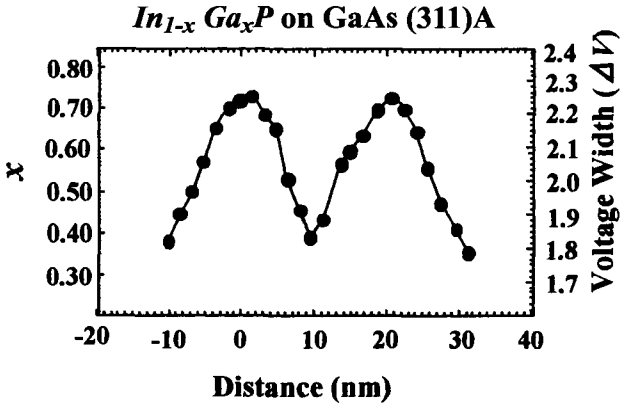
Grazing incidence small-angle X-ray scattering[26] and grazing incidence X-ray diffraction were used for structural characterization of InAs QDs. The grazing incidence technique probes the crystal surface up to depths of only a few nanometers. This strongly enhances the scattering intensity from the InAs QDs. It is revealed through this method that the lateral distribution of InAs QDs is anisotropic, and the most pronounced ordering of dot



(a)



(b)



(c)

Fig. 4 (a) STM image for QD structures self-formed $(GaP)_{1.5}(InP)_{1.88}$ SLs grown on the GaAs (311)A substrate, (b) dI/dV vs. V curves at different points, and (c) variation of the voltage width for $dI/dV=0$ in (b) along the [0 −1 1] direction. The left ordinate is related to x if the dot mean composition is represented by $In_{1-x}Ga_xP$. (From Ref. [20].) *(View this art in color at www.dekker.com.)*

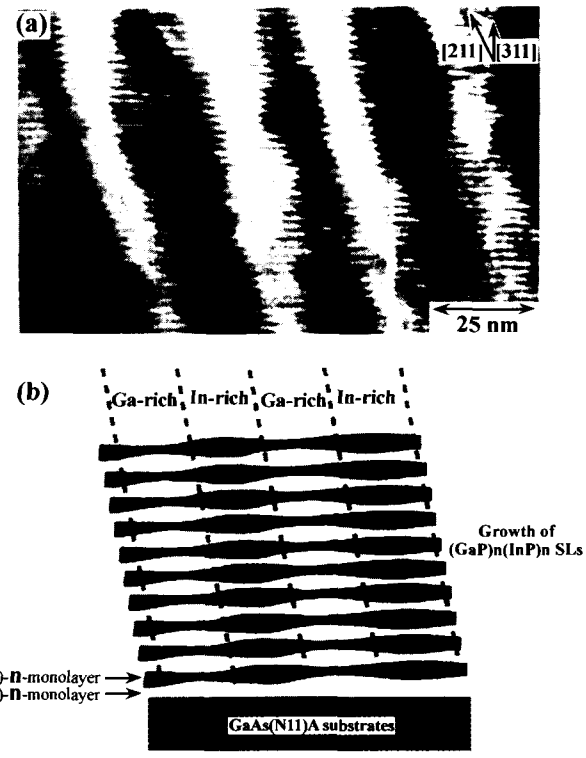


Fig. 5 (a) Cross-sectional TEM lattice image for the $(GaP)_{1.5}$ $(InP)_{1.88}$ sample grown on GaAs (311)A substrate, (b) schematic drawing illustrating the lateral composition modulation induced by the lateral thickness modulation. (From Ref. [3].)

distribution is in the [110] direction and that the dot shape is an octagonal-based truncated pyramid with {111} and {101} facet families.[27]

OPTICAL CHARACTERIZATION

In QDs, the electronic structures such as the energy level and the density of state greatly differ from those of bulk crystal. Therefore it is possible to characterize quantum dots by measuring QDs' optical properties, which reflect the electronic structures.

Photoluminescence (PL)

Ordinary PL peaks at low temperatures correspond to the energy difference between the lowest energy level in the conduction band of dots and the highest level in the valence band of the dots. Because these levels are a function of the size or radius of the dots, the peak position reflects the size of dots and the peak width reflects the size fluctuation of dots.

Fig. 7 shows an example of room-temperature PL spectra obtained by Songmuang et al.[28] for InAs dots

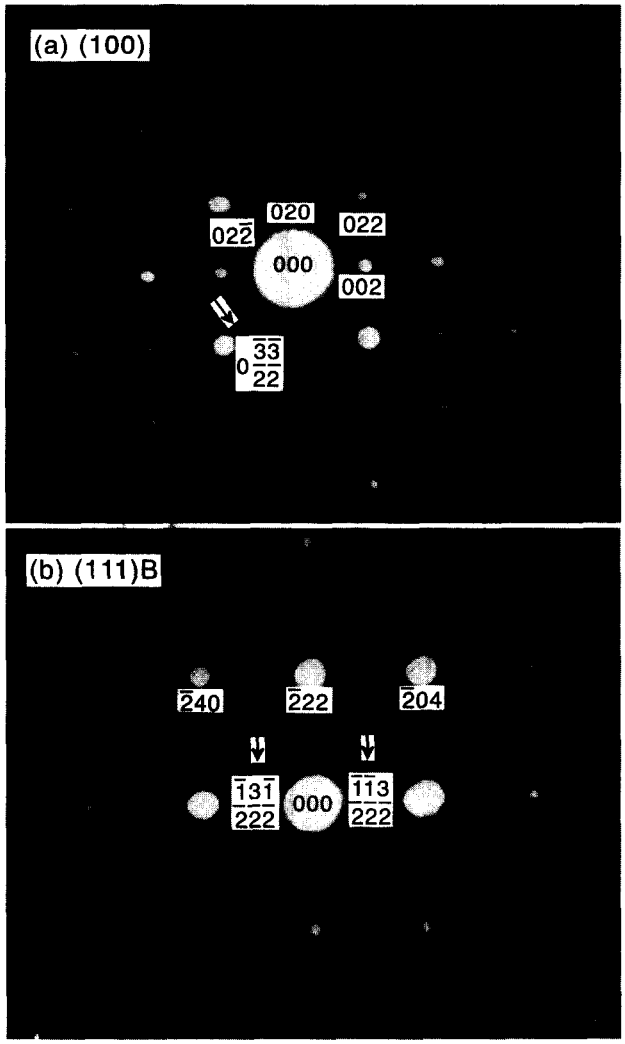


Fig. 6 TEM diffraction pattern for the (GaP)(InP) SLs grown on (a) GaAs (100) and (b) GaAs (111)B substrates. (From Ref. [25].)

fabricated through the S–K mode growth. The peak energy around 1 eV for InAs dots is very high, compared with the energy gap (0.36 eV) for InAs at room temperature. This shows the energy increase between the energy levels by dot formation. In this work, GaAs overgrowth on InAs dots was induced at two different temperatures. One is grown at the same level as the QD growth temperature [500°C; conventionally capped (CC)] and another at a lower temperature [470°C; low-temperature capped (LTC)]. In the lower spectra, the full-width at half maximum (FWHM) of the PL peak of CC dots is 51 meV, and that of LTC dots decreases to 26 meV. This implies that the size fluctuations decrease with decreasing overgrowth temperature.

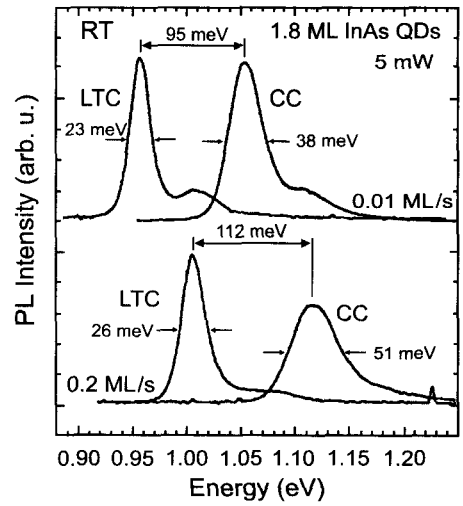


Fig. 7 PL spectra at room temperature for low-temperature capped (LTC) InAs dots and conventionally capped (CC) InAs dots by the SK mode growth method. (From Ref. [28].)

In the PL measurement of QDs fabricated by composition modulation, multilayer QDs, shown in Fig. 8, are used.[29] PL spectra at 77 K for $(InP)_{1.88}(GaP)_{1.5}$ and $(GaP)_{1.5}(InP)_{1.88}$ samples are shown in Fig. 9.[3] PL peak energy of the $(GaP)_{1.5}(InP)_{1.88}$ sample is higher than that of the $(InP)_{1.88}(GaP)_{1.5}$ sample by about 10%. This tendency corresponds to the difference in the lateral period of dots of two samples. The FWHM of PL peak of the $(InP)_{1.88}(GaP)_{1.5}$ sample is wider than that of the $(GaP)_{1.5}(InP)_{1.88}$ sample by about 10%, as shown in Fig. 9.

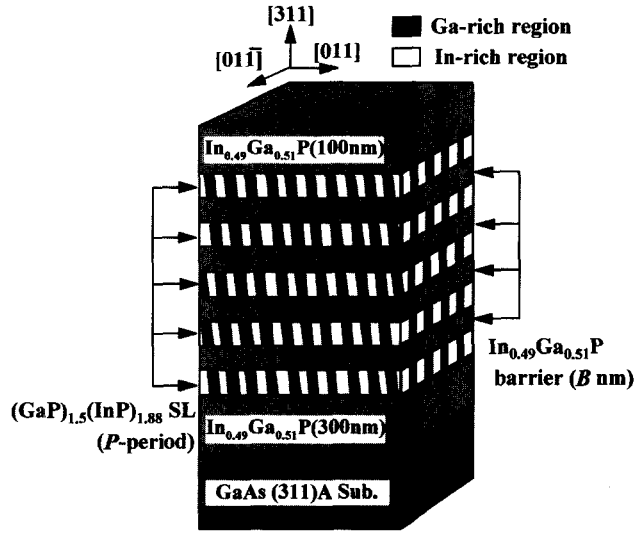


Fig. 8 Schematic drawing of five cycle multilayer QDs (MQD) formed in the $(GaP)_n(InP)_m$ SL/InGaP multilayer. (From Ref. [29].)

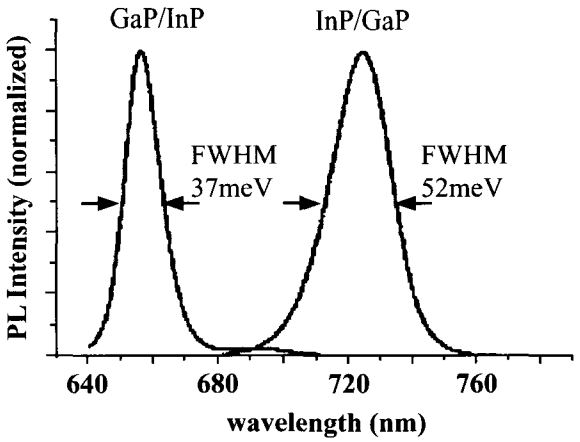


Fig. 9 PL spectra at 77 K for $(InP)_{1.5}(GaP)_{1.88}$ and $(GaP)_{1.5}(InP)_{1.88}$ samples. (From Ref. [3].)

The FWHM of the $(InP)_{1.88}(GaP)_{1.5}$ sample decreased nearly monotonically from 100 to 50 meV, when the temperature is lowered from 350 to 10 K. The FWHM of $(GaP)_{1.5}(InP)_{1.88}$ sample also decreases nearly monotonically down to 100 K, but it shows a maximum at $\sim$50 K. In this case, fluctuation of dot size is not to be related to the larger FWHM of the $(InP)_{1.88}(GaP)_{1.5}$ sample, because the deviation of period is nearly the same within the experimental error in both samples. The difference is probably a result of the difference in crystalline quality or the state of strain in the dot layers.

Optical properties of multilayer QDs made by composition modulation are investigated by changing the superlattice period and the InGaP barrier thickness. By decreasing the period, PL peak energy shifts toward higher energy because of the quantum size effect along the growth direction. The PL line width broadening with temperature is reduced by decreasing the period and the barrier thickness. This is attributed to the reduction of potential distribution among QDs and the enhancement of quantum confinement as a result of the vertical coupling effect between QDs. Very small temperature variation of PL peak energy is observed in multilayer QDs, which is attributed to the existence of the mutual strains in the multilayer QDs. A detailed discussion is given in another work.[30]

Time-Resolved Photoluminescence (TRPL)

TRPL is effective for the investigation of the nature of emission in quantum dots. TRPL measurements are performed with a fast streak camera in conjunction with a monochromator. Pulsed excitation is provided the frequency doubled beam of a mode-locked Al_2O_3:Ti laser.

An example of the result for multilayer QDs MQD5/20 (5 is the period of $(GaP)_{1.5}(InP)_{1.88}$ SL and 20 is the barrier thickness in nm) is shown in Fig. 10.[29] PL decay time strongly depends on emission energy. Fig. 11 shows the dependence of PL decay time on emission energy and temperature.[29] It ranges between 0.1 and 2.5 nsec. The decay time is shorter for higher energies. This dependence implies the existence of a tunneling process between QDs. In QDs, carriers in the higher energy levels of the smaller QDs can transfer to the lower energy levels in the adjacent larger QDs. Because carriers in the higher energy level have a large number of lower energy levels into which they can relax, the decay time becomes shorter for higher energies. Hence Fig. 11 reflects a variety of sizes of QDs and distances between QDs.

TRPL is also used to measure the tunneling time of carriers in coupled QDs. Takeuchi et al.[24] measured carrier tunneling times between vertically aligned double QDs using TRPL. The vertically aligned double QD structure consists of $In_{0.9}Ga_{0.1}As$ QDs, a GaAs barrier layer, and InAs QDs, which is made via the S–K mode growth. They revealed the dependence of the tunneling time on the barrier thickness.

Electroluminescence (EL)

EL shows another feature of the optical properties of QDs. We fabricated two types of light-emitting diodes (LED) using the composition modulation method. One is the sample MMQDs1, which has $(InGaP)_{40}(InAlP)_{40}$ cladding layers and $In_{0.49}Ga_{0.51}P$ barrier layers, and another is the sample MMQDs2, which has $(InGaP)_5(InAlP)_5$ clad-

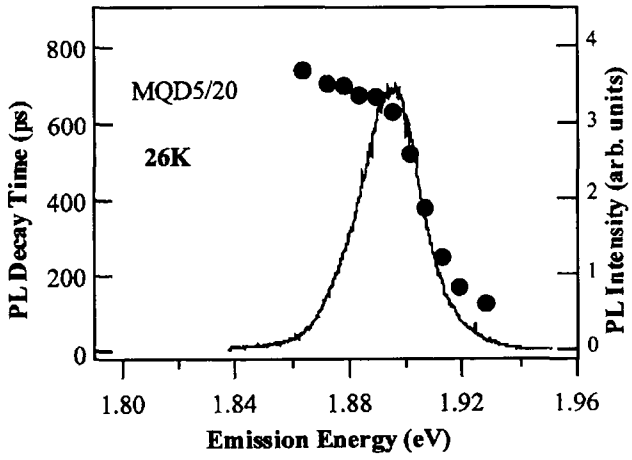


Fig. 10 Time-integrated PL spectrum and PL decay time as a function of monitored emission energy at 26 K for MQD5/20. (From Ref. [29].)

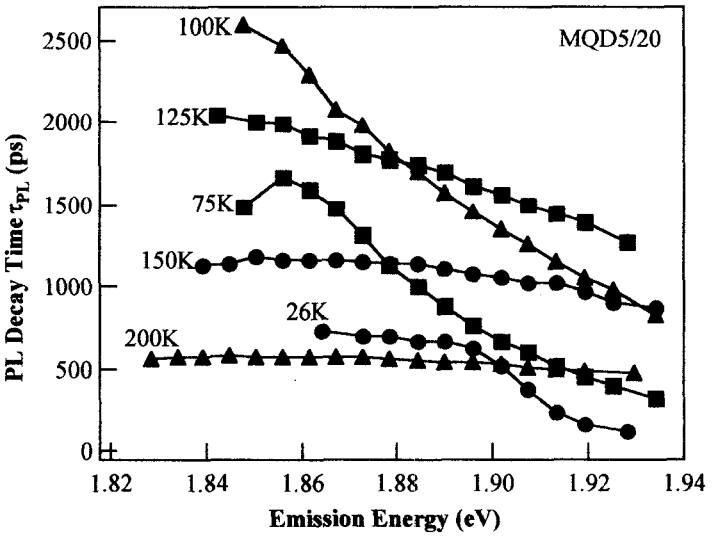


Fig. 11 Monitored emission energy dependence of PL decay time as a function of measured temperature. (From Ref. [29].)

ding and $(InGaP)_4(InAlP)_2$ barrier. EL spectra at room temperature for both samples are shown in Fig. 12.[31] MMQDs2 exhibits a narrower FWHM (48 meV) than that of MMQDs1. The integrated EL intensity of MMQDs2 is about 30 times higher than that of MMQDs1. This improvement results from the introduction of $(InGaP)_5$ $(InAlP)_5$ as cladding and $(InGaP)_4(InAlP)_2$ SL as barrier layers. Fig. 13 shows the temperature variations of integrated EL and PL intensities.[31] The temperature quenching of EL intensity is smaller for MMQDs2 than for MMQDs1. Because the effective band gap of the $(InGaP)_5(InAlP)_5$ SL is larger than that of $(InGaP)_{40}$-$(InAlP)_{40}$ SL, the reason for these temperature variations is considered to be a result of the suppression of carrier

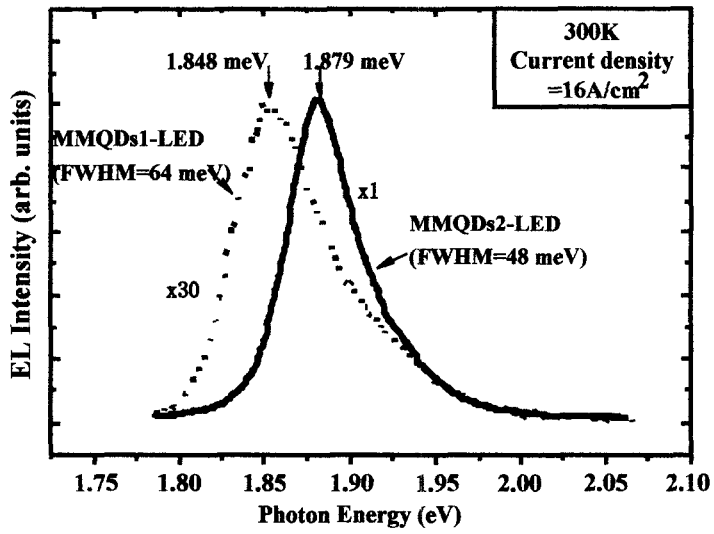


Fig. 12 EL spectra at room temperature. (From Ref. [31].)

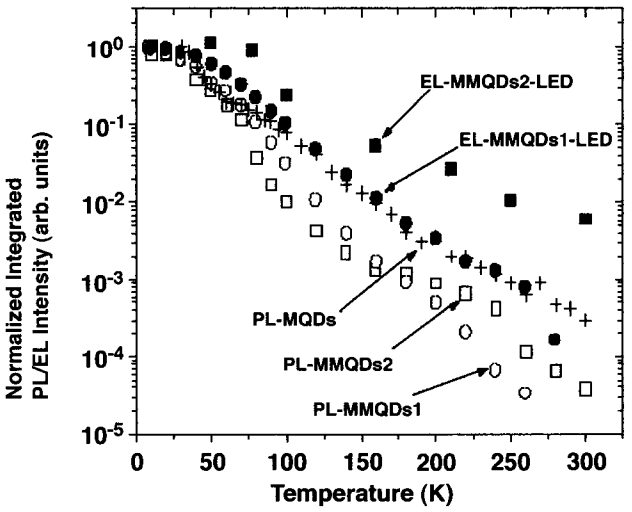


Fig. 13 Temperature variation of normalized EL and PL intensities. (From Ref. [31].)

flow from the QD region by the use of higher band gap cladding layers.

If the QDs are applied to devices, characterization via device performance is one of the very severe characterizations. The most popular application of self-formed QDs is semiconductor lasers. An example of InAs/InGaAs/GaAs QD lasers is introduced here. Kovsh et al.[32] made QDs via the S–K mode growth and fabricated 1.3-μm edge-emitting lasers. They simultaneously obtained threshold current density of 100 A/cm^2 and differential efficiency of 80% in the same device.

CONCLUSION

Structural and optical characterization of self-formed quantum dots was described, based mainly from our experience. Structures such as size, position, and distribution are well characterized by AFM, STM, STS, and TEM, and optical characterizations, capable of obtaining information on carriers and their energy states, are well carried out by PL measurements.

Most of the samples treated here were quantum dots fabricated via the composition modulation methods, and samples made by the S–K mode growth method are a little bit smaller in number. There are numerous reports concerning quantum dots produced by the S–K mode growth. For example, reports on In(Ga)As quantum dots are introduced in detail in Ref. [2]. For quantum dots of nitrides, the reader is referred to Ref. [9]. For further details on quantum dots using SiGe system, CdSe/ZnSe, and PbSe, please see Refs. [33–35] respectively.

The above-mentioned methods for structural and optical characterization are very effective and useful for various semiconductor materials.

REFERENCES

1. Bimberg, D.; Grundman, M.; Ledentsov, N.N. *Quantum Dot Heterostructures*; Wiley: New York, 1998.
2. *Self-Assembled InGaAs Quantum Dots, Semimetals and Semiconductors*; Sugawara, S., Ed.; Academic Press: New York, 1999; Vol. 60.
3. Gonda, S.; Asahi, H.; Mori, J.; Watanabe, D.; Matsuda, S.; Noh, J.H.; Fudeta, M.; Asami, K.; Seki, S.; Matsui, Y.; Tagawa, S. Structural and optical characterization of self-formed quantum dots. J. Electron. Mater. **2000**, *29*. 530–534.
4. Asahi, H. Self-organized quantum wires and dots in III–V semiconductors. Adv. Mater. **1997**, *9* (13), 1019–1026.
5. Leonard, D.; Krishnamruthy, M.; Reaves, C.M.; Denbarrs, S.P.; Petroff, P.M. Direct formation of quantum dots from uniform coherent islands of InGaAs on GaAs. Appl. Phys. Lett. **1993**, *63* (23), 3203–3205.
6. Ahopelto, J.; Yamaguchi, A.A.; Nishi, K.; Usui, A.; Sakaki, H. Nanoscale InP islands for quantum box structures by hydride vapor phase epitaxy. Jpn. J. Appl. Phys. **1993**, *32*, L32–L35.
7. Nabetani, Y.; Ishikawa, T.; Noda, S.; Sasaki, A. Initial growth stage and optical properties of a three dimensional InAs structure on GaAs. J. Appl. Phys. **1994**, *76*, 347–351.
8. Moisin, J.M.; Housy, F.; Barthe, F.; Leprince, L.; Andre, E.; Vatel, O. Self-organized growth of regular nanometer-scale InAs dots on GaAs. Appl. Phys. Lett. **1994**, *64*, 196–198.
9. Tachibana, K.; Someya, T.; Arakawa, Y. Nanometer-scale InGaN self-assembled quantum dots by metalorganic chemical vapor deposition. Appl. Phys. Lett. **1999**, *74* (3), 383–385.
10. Zhou, Y.K.; Asahi, H.; Asakura, J.; Okumura, J.; Asami, K.; Gonda, S. Magnetic and optical properties of Mn-including InAs dots grown by metalorganic molecular beam epitaxy. J. Cryst. Growth **2000**, *221*, 605–610.
11. Cheng, K.Y.; Hsieh, K.C.; Baillargeon, J.N. Formation of lateral quantum wells in vertical short-period superlattices by strain-induced lateral-layer ordering process. Appl. Phys. Lett. **1992**, *60* (23), 2892–2894.
12. Kim, S.J.; Asahi, H.; Takemoto, M.; Asami, K.; Takeuchi, M.; Gonda, S. Self-organized dot/columnar structure and quasi-perfect CuPt type ordering in $(GaP)_n(InP)_n$ superlattices grown on GaAs (N11) substrates by gas source molecular beam epitaxy. Jpn. J. Appl. Phys. **1996**, *35* (8), 4225–4231. Part 1.
13. Binnig, G.; Quate, C.F.; Gerber, Ch. Atomic force microscope. Phys. Rev. Lett. **1986**, *56* (9), 930–933.

14. Binnig, G.; Rohrer, H.; Gerber, Ch.; Weibel, E. 7×7 Reconstruction on Si(111) resolved unreal space. Phys. Rev. Lett. **1982**, *50* (2), 120–123.
15. Zhou, Y.K.; Asahi, H.; Asakura, J.; Okumura, T.; Tashima, T.; Sato, J.; Asami, K.; Gonda, S. Magnetic and optical properties of Mn-including InAs dots grown by MOMBE. Physica E **2000**, *7*, 1001–1005.
16. Leonard, D.; Fafard, S.; Pond, K.; Zhang, H.; Merz, J.L.; Petroff, P.M. Structural and optical properties of self-assembled InGaAs quantum dots. J. Vac. Sci. Technol. **1994**, *B12*, 2516–2520.
17. Ren, H.W.; Nishi, K.; Sugou, S.; Masumoto, Y. Size quantization in InAs/GaAs self-assembled quantum dots grown by gas-source molecular beam epitaxy. Jpn. J. Appl. Phys. **1998**, *37*, 1548–1551.
18. Martin, Y.; Wickramasinghe, H.K. Magnetic imaging by "force microscopy" with 1000 A resolution. Appl. Phys. Lett. **1987**, *50* (20), 1455–1457.
19. Tashima, T.; Asahi, H.; Sato, J.; Asami, K.; Zhou, Y.K.; Gonda, S. Formation of InAs dots including Mn atoms by metalorganic molecular beam epitaxy. Inst. Phys. Conf. Ser. **1999**, *162*, 481–486. Chapter 9.
20. Noh, J.H.; Asahi, H.; Kim, S.-J.; Gonda, S. Scanning tunneling microscopy study on self-formation process of quantum dot structures by the growth of GaP/InP short-period superlattices on GaAs(311)A substrate. Jpn. J. Appl. Phys. **1998**, *37* (6B), 3973–3975. Part 1.
21. Hasegawa, S.; Arakawa, K.; Tanaka, M.; Nakashima, H. Scanning tunneling spectroscopy study of InAs islanding on GaAs(001). J. Cryst. Growth **2001**, *227–228*, 1029–1033.
22. Joyce, P.B.; Krzyzewski, T.J.; Bell, G.R.; Jones, T.S.; Marik, S.; Childs, D.; Murray, R. Growth rate effects on the size, composition and optical properties of InAs/GaAs quantum dots grown by molecular beam epitaxy. J. Cryst. Growth **2001**, *227–228*, 1000–1004.
23. Ustinov, V.M.; Zhukov, A.E.; Maleev, N.A.; Kovsh, A.R.; Mikhrin, S.S.; Volovik, B.V.; Musikhin, Yu.G.; Shernyyakov, Yu.M.; Maximov, M.V.; Tsatsul'nikov, A.F.; Ledentsov, N.N.; Alferov, Zh.I.; Lott, J.A.; Bimbarg, D. 1.3 μm InAs/GaAs quantum dot lasers and VCSELs grown by molecular beam epitaxy. J. Cryst. Growth **2001**, *227–228*, 1155–1161.
24. Takeuchi, A.; Kuroda; Mase, K. Dynamics of carrier tunneling between vertically aligned double quantum dots. Phys. Rev. **2000**, *62*, 1568–1571.
25. Kim, S.J.; Asahi, H.; Takahashi, T.; Gonda, S. Substrate orientation dependence of lateral composition modulation in $(GaP)_n(InP)_n$ strained short period superlattices grown by gas source MBE. Inst. Phys. Conf. Ser. **1995**, *145*, 91–96. Chapter 2.
26. Schmidbauer, M.; Wiebach, Th.; Raidt, H.; Hanke, M.; Kohler, R.; Waura, H. Ordering of self-assembled $Si_{1-x}Ge_x$ islands studied by grazing incidence small-angle scattering and atomic force microscopy. Phys. Rev. **1998**, *B58*, 10523–10531.
27. Zhang, K.; Heyn, Ch.; Hansen, W.; Schmid, Th.; Felta, J. Structural characterization of self-assembled InAs quantum dots grown by MBE. J. Cryst. Growth **2001**, *227–228*, 1020–1024.
28. Songmuang, R.; Kiravittaya, S.; Sawadsaringkarn, M.; Panyakeow, S.; Schmidt, O.G. Photoluminescence investigation of low-temperature capped self-assembled InAs/GaAs quantum dots. J. Cryst. Growth **2003**, *251*, 166–171.
29. Fudeta, M.; Asahi, H.; Asami, K.; Nakamura, Y.; Kawakami, Y.; Noh, J.H.; Mori, J.; Watanabe, D.; Fujita, S.; Gonda, S. Time-resolved photoluminescence study of strain-induced quantum dots self-formed in GaP/InP short-period superlattices. Jpn. J. Appl. Phys. **1999**, *38* (9A/B), L1006–L1008. Part 2.
30. Kim, S.J.; Asahi, H.; Asami, K.; Takemoto, M.; Fudeta, M.; Gonda, S. Optiacal properties of quantum dots self-formed in GaP/InP short period superlattices grown on GaAs(N11) substrates. Appl. Surf. Sci. **1998**, *130–132*, 729–736.
31. Watanabe, D.; Asahi, H.; Noh, J.H.; Fudeta, M.; Mori, J.; Matsuda, S.; Asami, K.; Gonda, S. Improvement of optical properties of multilayer quantum dots self-formed in GaP/InP short-period superlattices on GaAs(311)A. Jpn. J. Appl. Phys. **2000**, *39* (7B), 4601–4603. Part 1.
32. Kovsh, A.R.; Maleev, N.A.; Zhukov, A.E.; Mikhrin, S.S.; Vasil'ev, A.P.; Semenova, E.A.; Shernyyakov, Yu.M.; Maximov, M.V.; Lifshits, D.A.; Ustinov, V.M.; Ledentsov, N.N.; Bimbarg, D.; Alferov, Zh.I. InAs/InGaAs/GaAs quantum dot lasers of 1.3 μm range with enhanced optical gain. J. Cryst. Growth **2003**, *251*, 729–736.
33. Jin, G.; Wan, J.; Luo, Y.H.; Liu, J.L.; Wang, K.L. Uniform and ordered self-assembled Ge dots on patterned Si substrates with selectively epitaxial growth technique. J. Cryst. Growth **2001**, *227–228*, 1100–1105.
34. Maehashi, K.; Yasui, N.; Ota, T.; Noma, T.; Murase, Y.; Nakashima, H. Structural and optical properties of CdSe/ZnSe self-organized quantum dots. J. Cryst. Growth **2001**, *227–228*, 1116–1120.
35. Springhotz, G.; Pinczolits, M.; Bauer, G.; Knag, H.H.; Salamanca-Riba, L. Phase diagram of lateral and vertical ordering in self-organized PbSe quantum dots grown by MBE. J. Cryst. Growth **2001**, *227–228*, 1126–1131.

Quantum Dots, Semiconductor: Atomic Ordering over Time

Q

Peter Moeck
Portland State University, Portland, Oregon, U.S.A.

INTRODUCTION

This encyclopedia entry consists of two parts. The first part gives a brief overview of epitaxially grown semiconductor quantum dots. This kind of quantum dot is referred to as ordinarily strained random semiconductor alloy quantum dot throughout this review. The objective of the first part is to set the scene for the following part.

The main part of this encyclopedia entry deals with experimental evidence in favor of the existence of long-range atomic order within certain III–V and II–VI compound semiconductor Stranski–Krastanow grown quantum dots a few years after the growth. The objective of the second part of this entry is to review this particular field of inquiry. With necessity, only the most general experimental results can be discussed. A simple thermodynamic model, which, in principle, allows for structural transitions from ordinarily strained random semiconductor alloy quantum dots with the structural prototype of the surrounding matrix to crystallographic superlattices (i.e., long-range atomically ordered quantum dots), is also briefly mentioned in the second part.

EPITAXIALLY GROWN ORDINARILY STRAINED RANDOM SEMICONDUCTOR ALLOY QUANTUM DOTS

Epitaxial semiconductor quantum dots (QDs) are expected to lead to "paradigm changes in semiconductor physics".[1] As reviewed in recent textbooks,[2,3] improved optoelectronic devices (such as lasers[4,5] and mid-infrared detectors[6]). and novel nanoelectronics concepts which depend on single electron transport, tunneling, or interactions (e.g., quantum cellular automata[7]) have either been realized or may be realized as the field matures.

For applications in optoelectronic devices, an entity needs to fulfill three conditions simultaneously to be considered a QD.[4] The entity needs to possess an appropriate size in all three dimensions (i) and also needs to be a semiconductor with a smaller band gap that is embedded in a semiconductor matrix with a larger band gap (ii). No detrimental structural defects (such as dislocations, iii), which lead to nonradiative recombination centers, are allowed to exist in the entity.

Note here that there is no restriction on the structural prototype of the entity. It may either be of the same type as that of the matrix or it may possess a distinctly different structural prototype to that of the matrix. Structural prototypes of an epitaxially grown entity that differ from that of the matrix may be realized over time by atomic rearrangements within the initially more or less random semiconductor alloy entity. In short, atomically ordered entities that may arise from such processes will have to be considered as constituting QDs according to the definition above when conditions i to iii are met simultaneously.

Analogous to the well-known particle in a box model, the simplest quantum mechanical model describes a QD as a three-dimensional (3-D) finite potential-barrier trap for an electron–hole pair. When the size of this trap is in all three dimensions smaller than the bulk exciton Bohr radius (order of magnitude 10 nm), discrete energy levels for the electron and hole arise from the spatial confinement of these particles. The energy separation of the electron levels is, for example, for epitaxially grown (In,Ga)As QDs, of the order of magnitude 100 meV,[1,8] i.e., significantly larger than the thermal energy at room temperature ($kT \approx 26$ meV). It is these discrete energy states of an electron in a QD that are exploited in devices that work at room temperature.

As in the particle in a box model, the spacing of the energy levels in QDs depends sensitively[8] on a range of parameters of the matter wave trap. For a QD, these parameters are first and foremost its size, shape, net chemical composition and spatial chemical composition distribution, net lattice mismatch strain and spatial strain distribution, crystallographic phase, and orientation with respect to the surrounding matrix. Because a very large number of QDs are simultaneously employed in optoelectronic devices, it is obvious that variations in the parameters of the individual matter wave traps are to be as small as possible to circumvent undue broadening of the discrete electron energy levels on which the operation of

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013806

the devices relies. Although addressing this uniformity issue of QDs by optimized growth and postgrowth annealing processes is still a scientific challenge, optimizations of the basic epitaxial growth processes have been achieved in recent years on the basis of structural and spectroscopic analyses.

Currently most popular with the scientific community are QDs that are grown in the Stranski–Krastanow mode and its variants (such as embedding nominal submonolayers and allowing them to form 3-D entities within a 2-D wetting layer).[a] Either molecular beam epitaxy (MBE) or metal organic chemical vapor deposition (MOCVD) is typically employed.[2,3] Because these QDs grow epitaxially, they possess per definition of the term epitaxy the same structural prototype as the employed substrates and surrounding matrices. Epitaxial growth of QDs typically results in a compressive lattice mismatch strain in the QDs of a few percent as the QDs need to have a smaller band gap than the surrounding matrix and smaller band gaps are typically obtained for solid substitutional solutions with larger lattice constants.

The nature of the Stranski–Krastanow 2-D to 3-D transition in (In,Ga)As has recently been identified, and it is believed that QD of other semiconductor systems grows in the same basic manner.[11–13] Because of the nature of this transition, the QDs and their predecessor islands are always alloyed, should have a more or less random distribution of the mixed cations and/or anions over their respective sublattices in the structural prototype of the employed substrate, and should have a similar net chemical composition regardless of whether nominally pure substances or substitutional solid solutions are deposited.

Structural studies by numerous investigators employing conventional transmission electron microscopy (TEM), high-resolution phase contrast TEM (HRTEM), and atomic resolution Z-contrast scanning TEM (Z-STEM) of a wide range of Stranski–Krastanow as-grown III–V and II–VI QD and QD predecessor island samples showed (or at least reported) no evidence of atomic ordering in as-grown structures, e.g., Refs. [11–21] and references therein. In addition, it is known that the predecessor islands and QDs have a spatially nonuniform chemical composition distribution with a core that is rich of the component with the smaller band gap (and typically larger lattice constant), i.e., a spatially zonal chemical composition distribution.[8,14,15] Significant morphological transformations[16] and changes in both the net chemical composition and spatial composition distribution[14] of the predecessor island structures take place as a result of the QD embedding process.

Such Stranski–Krastanow mode grown QDs are referred to as "ordinarily strained" in the remainder of this encyclopedia entry because they show pronounced strain fields (either so-called "black-white" or "coffee-bean" contrasts) in conventional diffraction contrast TEM micrographs, e.g., Fig. 1a.[17–19] The author of this encyclopedia entry will use the quantifier "random semiconductor alloy" synonymously to refer to ordinarily strained QDs. Occasionally, the author will address these

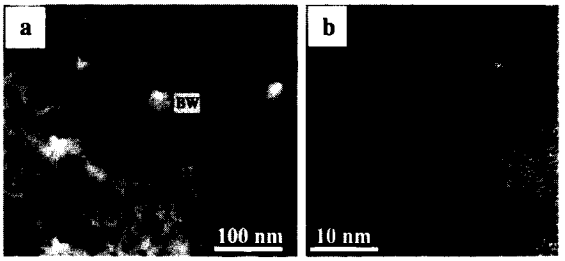


Fig. 1 (a) Ordinarily strained (In,Ga)Sb quantum dots in GaSb matrix shortly after the growth, showing strain contrast in conventional diffraction contrast TEM images, [001] plan view; the marker "BW" adjacent to a QD stands for "black-white contrast." (b) (Partly) relaxed (In,Ga)Sb agglomerate in GaSb matrix a few years after the growth, showing $\pm(002)$ and $\pm(2\bar{2}0)$ moiré fringes, $\langle 110 \rangle$ cross-section HRTEM image, Fourier transform power spectrum as insert, modified after Ref. [27] and reproduced with permission of the International Society for Optical Engineering (SPIE). The shape of the agglomerate is that of an oblate spheroid/ellipsoid and is believed to result from strain-driven atomic rearrangements in the sphalerite structure. It is well known that if QD predecessor islands grow beyond a certain size/volume limit, the lattice mismatch strain is relaxed to a significant amount by the introduction of misfit dislocation.[17,18,21] While the diameter of the quantum dots in (a) is about 10 nm, i.e., below the above-mentioned size limit, the approximate width of the agglomerate in (b) is 47 nm and its height is approximately 28 nm. These latter dimensions result in a volume that is conceivably above the volume limit for lattice mismatch strain relaxation.

[a]What exactly Stranski–Krastanow growth means in semiconductor epitaxy is seldom defined. The classical texts and undergraduate textbooks (e.g., Ref. [9]) define this mode simply as one of the three fundamental modes of heteroepitaxial growth (besides the Frank–Van der Merwe and the Volmer–Weber growth mode). Stranski–Krastanow growth is typically described as comprising a sequence in epitaxial growth where first, a 2-D wetting layer is formed, and second, 3-D islands are formed on top of this wetting layer. No particular mechanism is associated with the 2-D to 3-D transition per definition, and the original paper Ref. [10] deals only with the energetics of the epitaxial growth of lattice-matched ionic crystals. With explicit mechanisms not being connected to this growth mode, one is free to classify any epitaxial growth where there is a 2-D to 3-D transition of the deposit as Stranski–Krastanow growth. Whether the 3-D entities are formed on top of or within a 2-D wetting layer is thus quite unimportant. Embedding nominal submonolayers of a smaller band gap within a matrix of a larger band gap and allowing them to form 3-D entities within a 2-D wetting layer is therefore just a variant of Stranski–Krastanow growth in semiconductor epitaxy.

QDs with both quantifiers, i.e., as "ordinarily strained random semiconductor alloy QDs," to emphasize the differences to atomically ordered QDs. Where appropriate, this author will also specify the semiconductor in the general specifier "semiconductor alloy."

Structural analyses of QDs are mainly performed by transmission electron microscopy and near surface-sensitive X-ray scattering techniques. Only very few of such investigations, however, address the long-term structural and morphological stability of QDs at room temperature over a time period of a few years. Such TEM investigations are therefore reviewed in the second part of this encyclopedia entry.

ATOMIC ORDER WITHIN EPITAXIAL SEMICONDUCTOR QUANTUM DOTS

As mentioned above, reviewing experimental results on morphological transformations[21] and structural transition[22–32] in III–V and II–VI compound semiconductor Stranski–Krastanow grown QDs will be the subject of this part of this encyclopedia entry. Because a total of eight original papers are to be reviewed in this part, only the most general result, i.e., the fact that long-range atomic order is present in certain epitaxial compound semiconductor QD systems after the respective samples have been stored at room temperature for a few years, will be discussed in some detail. A variety of TEM images from different QD systems will be shown to demonstrate different kinds of this atomic order. This multitude of TEM images serves the purpose to show the readers (and fellow electron microscopists) the different appearances long-range atomic ordering may take in Stranski–Krastanow grown QDs of different materials systems. Occasionally, the figure captions contain details that complement the text.

The respective structural analyses at the University of Illinois at Chicago, Portland State University, and the University of Oxford, U.K., employed TEM in both the parallel illumination and scanning probe (STEM) mode.[21–32] A JEOL JEM-2010F STEM/TEM, a JEOL JEM-3010 TEM, a Philips CM20 TEM/STEM, and a FEI Tecnai G^2 F20 TEM/STEM were used. Parallel illumination TEM utilized conventional diffraction contrast (CTEM), selected area electron diffraction (SAED), and high-resolution phase contrast imaging (HRTEM). Atomic resolution Z-contrast (also known as high-angle annular dark field) imaging in the scanning probe mode (Z-STEM) proved to be especially useful for these investigations because the effects of strain fields in and around QDs and interference effects such as the formation of moiré fringes because of double diffraction are negligible.[30–32] Transmission electron microscopy specimen preparation artifacts were excluded from the analyses by employing both ion milling and chemical thinning to electron transparency. Only those structures that are not an artifact of either of these two physically different specimen preparation processes were likely to be present in both kinds of specimen.[b]

Complementary photoluminescence (PL) spectroscopy analyses were performed at the Physical-technical Federal Institute Braunschweig/Germany on certain II–VI QD structures at a temperature of 6 K. Ar^+ laser light utilizing both the 488-nm line and ultraviolet light of around 350-nm wavelength was used at varying power densities for the excitation of the spectra. The PL emission light was dispersed in a 1-m-long single-grating spectrometer and detected with a cooled (In,Ga)As photomultiplier tube using conventional lock-in techniques.

Although there have been many investigations, there is still some controversy about the actual shape of both the islands that form the QDs when overgrown by a capping layer and of the ordinarily strained QDs themselves. Transmission electron microscopy observations on the shape of ordinarily strained QDs a few years after the growth resulted for a III–V system[21] (Fig. 1b) and a II–VI system,[20,30–32] in shapes that are oblate spheroids which may have different diameters along $\pm[110]$ and $\pm[\bar{1}10]$, making them, in effect, ellipsoids.[16] This shape is known as a possible equilibrium shape of small precipitates that are elastically strained.[35] Of each spheroid/ellipsoid, about one-half was located above the level of the original wetting layer and the other half was located below this level[20,21,30–32] (Fig. 1b).

[b]The close proximity of three different atomically ordered (Cd,Mn,Zn)Se agglomerates labeled by markers "1," "2," and "3" in Fig. 5a can only be interpreted as not being an artifact of the STEM imaging procedure because it is quite incomprehensible how one and the same probing electron beam may have caused three different modes of atomic rearrangements so close to each other. On a similar token, not only atomic ordering, but also phase separation were observed in the same TEM specimen (Fig. 6a–c), and it does not seem likely that the same probing electron beam may have triggered different kinds of atomic rearrangements in the same specimen. On the other hand, thermodynamics allows for the coexistence of atomically ordered and phase-separated entities in many compound semiconductor alloys. Ref. [33] Note also the large size of the double period $\pm(1\bar{1}0)$ agglomerates, i.e., order of magnitude 200 nm (Fig. 4c and d). These entities seem to be by far too large to be caused by the probing electron beam, which had a diameter on the order of magnitude 0.2 nm only. The results of Ref. [34] on In agglomeration into QDs within (In,Ga)N quantum wells because of parallel illumination and scanning probe electron beam exposures do not reveal any changes at the structural prototype level. Such structural prototype transitions may, however, have occurred in the sample of Figs. 4c, d, and 5a. One can thus conclude that these entities are definitely not an artifact of the probing electron beam. Because it is quite impossible to produce three different atomically ordered (Cd,Mn,Zn)Se agglomerates in close proximity (Fig. 5a) by means of TEM specimen preparation procedures, such artifacts can also be excluded from the analyses.

The moiré fringes in Fig. 1b indicate that the (In,Ga)Sb entity possesses its own (partly) relaxed lattice constant, i.e., no longer (fully) strained. Because it must contain dislocations that mediate the relaxed state (i.e., violate the condition iii given in the first part of this entry), it cannot be considered to be a QD and is addressed as an "agglomerate." The degree of relaxation can only be estimated as the chemical composition of the agglomerate is unknown. This author assumes that the degree of relaxation is less than unity because it is well known that complete relaxation of strained semiconductor structures is rarely achieved. Because numerous STEM and TEM images of QD predecessor islands and QDs by other authors have shown rather different shapes to those oblate spheroids/ellipsoids that were reported in Refs. [20,21,30–32], this author assumes that there are lattice mismatch strain energy-driven morphological transformations over time. As the (In,Ga)Sb agglomerate in Fig. 1b is obviously dislocated but possesses the shape of an oblate spheroid/ellipsoid, even small lattice mismatch strains may have the capacity to cause morphological transformations over time.

Now if one accepts that morphological transformation of QDs may take place over time, the conceptional barrier to accepting that structural transitions in QDs may also take place over a sufficiently long time is lowered. The amount of elastic mismatch strain energy that is stored in ordinarily strained QDs and (partly) relaxed agglomerates may, besides the net chemical composition, be a determining factor for the type of transformation. Either morphological, structural, or both structural and morphological transformations may thus take place over time. If this strain energy is comparably small (and/or the net chemical composition is unfavorable), as possibly in the case of (partly) relaxed (In,Ga)Sb agglomerates, e.g., Fig. 1b, only morphological transformations may be possible. If, on the other hand, this strain energy is comparably large (and/or the net chemical composition favorable), as possibly in the case of fully strained (In,Ga)Sb QDs, e.g., Fig. 1a, structural transitions may be possible.

Continuing with the (In,Ga)Sb QDs in GaSb matrix system, the results of two CTEM investigations[18,19] on the same TEM specimens that were separated by approximately 2 years in time suggest that the number density of ordinarily strained QDs declined over time. This observations could not be explained in 1999, but HRTEM and atomic resolution Z-STEM investigations that were performed on the same specimen about 1 to 2 years later revealed the presence of atomically ordered QDs in these specimen (Fig. 2a–c). If one assumes that ordinarily strained QDs transform structurally into atomically ordered QDs over time while the specimens are stored at room temperature, there seems to be a simple explanation for this inferred decline in the number density of ordinarily strained QDs. In both studies, the ordinarily strained quantum dots were detected by their pronounced strain field contrasts, e.g., Fig. 1a, and only moderate magnifications were employed to cover large specimen areas in a Philips CM 20 (that was not capable of high-resolution phase contrast imaging of medium lattice constant semiconductors). Note here that atomically ordered QDs, e.g., Fig. 2a–c, would, under these experimental conditions,[18,19] be effectively undetectable as they are negligibly strained and do not show pronounced strain field contrasts.

It is also noteworthy that short-range order, but no long-range order, was observed in (In,Ga)Sb QDs in GaSb matrix specimen by means of SAED approximately 14 days after the growth.[17,26] Similarly to this III–V QD system, both short-range order and beginning long-range atomic order were observed by means of HRTEM in

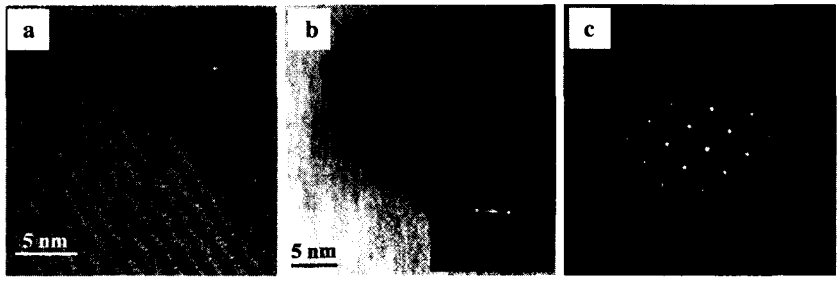


Fig. 2 Long-range atomically ordered (In,Ga)Sb quantum dots in GaSb matrix a few years after the growth, Fourier transform power spectra as inserts. (a) ⟨110⟩ Cross-section Z-contrast STEM image, showing a QD with atomic order in every fourth ±(002) plane. (b) [001] Plan-view HRTEM image which was recorded at 500°C after a thermal treatment in the electron microscope at temperatures of several hundred degrees centigrade for several hours, suggesting that the atomically ordered structure of this QD rather than its original sphalerite prototype structure is thermodynamically stable. (c) [001] Plan-view SAED pattern to (b) showing the same superlattice spots as the Fourier transform power spectrum insert in (b). This diffraction pattern demonstrates clearly that atomically ordered QDs such as shown in (b) are to be distributed over large areas of the specimen. (a) and (b) were modified after Ref. [27] and reproduced with permission of SPIE.

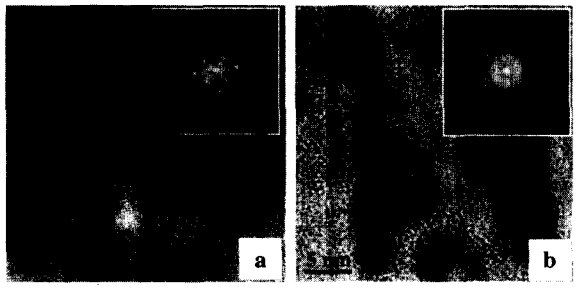


Fig. 3 High-resolution TEM [001] plan-view images of atomically ordered (Cd,Zn)Se quantum dots in ZnSe matrix a few years after the growth, showing (a) long-range and (b) short-range atomic order in the same TEM specimen. The insert Fourier transforms power spectra show for (a) superlattice spots and for (b) both a diffuse and a spotty ring. None of this atomic order seems to have existed at the time of a much earlier atomic resolution Z-STEM investigation. (From Refs. [20,30–32].)

(Cd,Zn)Se QDs in ZnSe matrix samples (Fig. 3a and b) about 3 years after the growth.[26] Samples from the same growth run were much earlier analyzed by means of atomic resolution Z-STEM, but no atomic order seemed to have been present at that earlier time.[20]

These HRTEM and atomic resolution Z-STEM results from both a III–V and a II–VI system indicate that semiconductor QDs when initially grown are mostly unordered, but during subsequent room temperature storage over a few years (i.e., a long-time low-temperature annealing process), they progressively undergo short-range ordering that eventually goes over into long-range ordering.

Long-range atomically ordered QDs and entities that were probably too large to be considered as a QD (i.e., that probably violated condition i given in the first part of this entry) were also observed in MBE-grown (Cd,Mn,Zn)Se QDs in (Mn,Zn)Se matrix (Figs. 4a–f and 5a).

As the Fourier transform power spectra (Fig. 4e and f) show, the atomic arrangement in the large (Cd,Mn,Zn)Se agglomerate in Fig. 4c and d is a result of atomic ordering of the Cd and (Zn,Mn) cations in every second $\pm(1\bar{1}0)$ plane. Using the shortest vectors in Fig. 4e as a base (for the sake of coming up with a notation only and assuming that the agglomerate belongs at least to the orthorhombic crystal system), the orientation relationship between the agglomerate and the (Mn,Zn)Se matrix becomes $(020)_{(Cd,Mn,Zn)Se} \| (2\bar{2}0)_{(Mn,Zn)Se}$, $(110)_{(Cd,Mn,Zn)Se} \| (\bar{1}13)_{(Mn,Zn)Se}$, and $[001]_{(Cd,Mn,Zn)Se} \| [110]_{(Mn,Zn)Se}$ (whereby $\|$ symbolizes that both sets of reciprocal lattice vectors are within the accuracy of the Fourier transform power spectra parallel and of equal length, and that the direct lattice vectors are parallel). This orientation

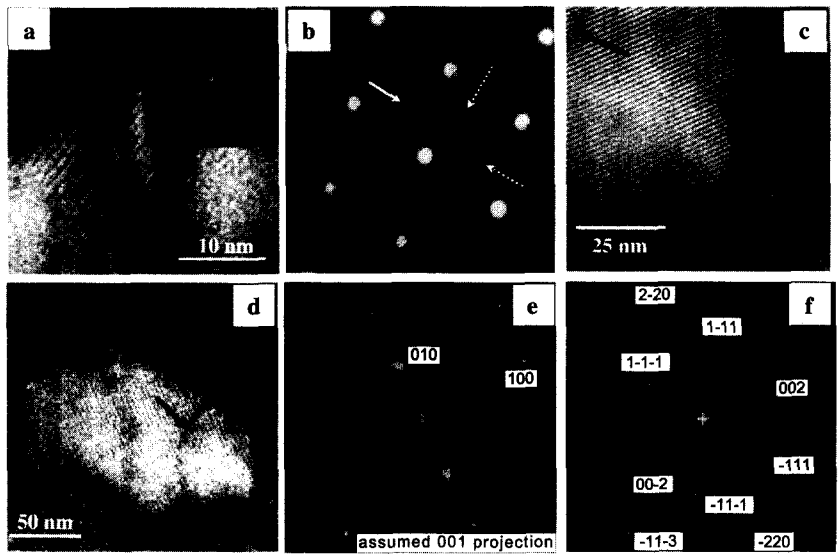


Fig. 4 TEM/STEM images of long-range atomically ordered (Cd,Mn,Zn)Se quantum dots and agglomerates in (Mn,Zn)Se matrix a few years after the growth; (a), (b), (d), (e), and (f) modified after Ref. [27] and reproduced with permission of SPIE. (a) ⟨110⟩ Cross-section HRTEM image of a (Cd,Mn,Zn)Se QD in (Mn,Zn)Se matrix, showing cation ordering in every second $\pm(1\bar{1}1)$ plane, i.e., one variant of so-called CuPt-type order, Fourier transform power spectrum as insert. (b) ⟨110⟩ Cross-section SAED pattern of (Cd,Mn,Zn)Se QDs in (Mn,Zn)Se matrix, showing $\pm\frac{1}{3}(2\bar{2}0)$, full line arrow, $\pm\frac{1}{3}(\bar{1}13)$, and $\pm\frac{1}{3}(1\bar{1}3)$, broken line arrows, superstructure spots. (c) ⟨110⟩ Cross-section atomic resolution Z-STEM image of a part of a large (Cd,Mn,Zn)Se agglomerate in (Mn,Zn)Se matrix, showing atomic order in every second $\pm(1\bar{1}0)$ plane, the black arrow points to an antiphase boundary. (d) Lower magnification Z-STEM image to (c), the black arrow points to the same antiphase boundary. (e) Fourier transform power spectrum of the large (Cd,Mn,Zn)Se agglomerate in (c). (f) Fourier transform power spectrum of the (Mn,Zn)Se matrix in (c).

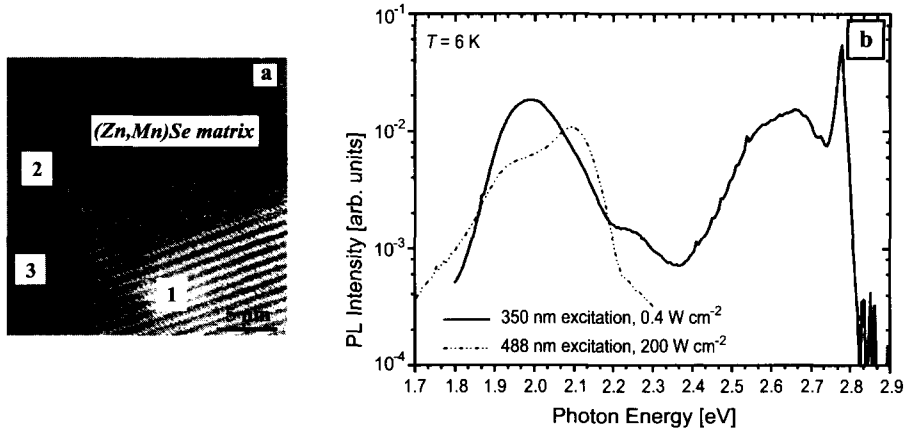


Fig. 5 (a) ⟨110⟩ Cross-section atomic resolution Z-contrast STEM image of (Cd,Mn,Zn)Se entities in a (Mn,Zn)Se matrix, a few years after the growth; modified after Ref. [27] and reproduced with permission of SPIE. Note that there are three different long-range atomically ordered entities, labeled as "1," "2," and "3" in close proximity; this remarkable fact is discussed in footnote c as proof that the observed atomic order in QDs is not a result of specimen preparation or probing electron beam artifacts. (b) Photoluminescence spectra of the same (Cd,Mn,Zn)Se entities in a (Mn,Zn)Se matrix. While the nonresonant PL is given by a full line. the resonant PL is given by a dotted line. As the area from which the PL spectra were taken is about 6 orders of magnitude larger than the area of a typical TEM image, this figure proves that atomically ordered entities must be present with a larger number density; courtesy of K. Pierz.

relationship is obviously lattice mismatch strain (and elastic mismatch strain energy) minimizing.

In the same TEM/STEM specimen, there were actually several more large agglomerates which possessed the same lattice mismatch strain energy-minimizing orientation relationship as shown in Fig. 4e and f. From a discussion of the nominal growth condition of this sample,[c] the large size of these agglomerates cannot be explained. Some of these agglomerates had diameters of up to approximately 250 nm and were also free of detrimental structural defects such as dislocations. The author of this encyclopedia entry believes that it is the elastic mismatch strain energy minimization effect of the particular orientation relationship these agglomerates possess with the (Mn,Zn)Se matrix which allows them to be so large while retaining their ideal translational symmetry. Taking the discussion in footnote c into account, the large size of the agglomerates may thus be explained by both the particular elastic mismatch strain energy-minimizing orientation relationship and atomic rearrangements over time. From the thermodynamics of small misfitting precipitates,[35] one can simply argue that smaller entities of the same atomically ordered structure and orientation relationship will not contain misfit dislocations either, i.e., will constitute atomically ordered QDs in this materials system.

Having identified large double-period $\pm(1\bar{1}0)$ ordered (Cd,Mn,Zn)Se agglomerates as by far the dominant kind of structural entity per unit volume of the TEM specimens, it is now interesting to analyze the PL spectra of this sample (Fig. 5b). The ZnSe band gap being at 6 K approximately 2.8 eV, excitation with 350 nm light can be considered as nonresonant above this band gap and is expected to reveal most radiatively active structures inside the samples. Light excitation of 488 nm, on the other hand, can be considered as resonant excitation below the ZnSe band gap, is typically of orders of magnitude weaker than nonresonant excitation over the band gap, and is selective of specific structures.

It was suggested that the large and rather broad PL peak at 2 eV in Fig. 5b, which arose when a sample was excited with laser light of approximately 350 nm, is a result of larger and smaller double-period $\pm(1\bar{1}0)$ agglomerates, as shown in Fig. 4c and d and labeled "1" in Fig. 5a.[28] The 2.1-eV PL peak (Fig. 5b), which arose in addition to the 2-eV peak, when the sample was excited with laser light of 488 nm, may be a result of either small double-period $\pm(1\bar{1}1)$ atomically ordered QDs (Fig. 4a), antiphase boundaries such as marked by an arrow in Fig. 4c and d, or other long-range atomically ordered entities such as labeled "2" and "3" in Fig. 5a.

[c]It is interesting to contrast the large size of these double period $\pm(1\bar{1}0)$ atomically ordered agglomerates with the nominal as grown sample structure. Starting with a ZnSe/GaAs pseudosubstrate, there should nominally be a multilayer structure of eight sequences of 10 monolayers (2.83 nm) of $Zn_{0.9}Mn_{0.1}Se$ cladding layer and 0.3 monolayers (0.09 nm) CdSe sheet, capped by a further 10 monolayers of $Zn_{0.9}Mn_{0.1}Se$ and 50 nm of ZnSe. The agglomerate in Fig. 4c and d being approximately 100 nm thick is thus thicker than the whole nominal epitaxial multiquantum well structure plus the two capping layers. It seems therefore highly improbable that such a large agglomerate could have originated directly during the growth of the QD structure a few years ago. Note that "remains" of this nominal as-grown structure were actually observed by means of HRTEM. Ref. [23,28].

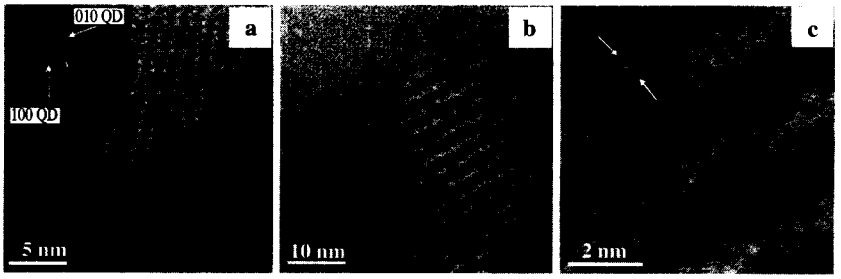


Fig. 6 [001] Plan view atomic resolution Z-contrast STEM images of long-range atomically ordered and phase-separated In(As,Sb) quantum dots in InAs matrix, a few years after the growth; modified after Ref. [27] and reproduced with permission of SPIE. (a) Atomically ordered QD that possesses a lattice mismatch strain energy minimizing orientation relationship with the matrix which can be inferred from the insert Fourier transform power spectrum. (b) Periodic compositional modulation that may have arisen from spinodal decomposition. (c) Same as (b) but in a higher magnification, showing atomic steps in the compositionally modulated entities that are compatible with the two observed additional periodicities in the insert Fourier transform power spectrum. The power spectrum clarifies that the modulation periodicity is not a result of a moiré effect because there are actually two additional periodicities, marked by arrows. These two periodicities add up to a single compositional modulation parallel to ±[110] with a periodicity of 0.4 times the length of the vector (110) and may be explained as a lattice mismatch strain response. In addition, Z-contrast imaging in STEM is known to be essentially incoherent so that double diffraction effects are negligible.

From Fig. 5b, it can be concluded that the long-range atomically ordered entities must be distributed throughout large regions of the sample because the PL spectra were collected from areas that were about 6 orders of magnitude larger than the areas that were investigated by means of TEM (Fig. 5a). A brief review of the literature showed that low-energy peaks at approximately 2 and 2.1 eV were also observed in this QD system in comparable nominal structures by several other authors.[28] None of these other authors, however, supported their assignment of these PL peaks with structural characterizations.[d]

The observations on both III–V and II–VI QD systems with mixed cations are thus consistent with the working hypothesis that atomic rearrangements that lead to lower-energy structures take place over a sufficiently long time. Some growth-induced short-range order may constitute additional "seeds" from which long-range ordered QDs could grow over time. It is interesting that in a III–V quantum dot system with mixed anions, i.e., In(As,Sb) QDs in InAs matrix (Fig. 6a–c), atomically ordered QDs and phase-separated entities were also observed.

Labeling the two shortest vectors in the Fourier transform power spectrum of Fig. 6a as a base of the reciprocal lattice of the atomically ordered In(As,Sb) QD (for the sake of coming up with a notation only), the insert of this figure shows that the ±(210) vectors of this QD are almost at the same position in reciprocal space as the ±(110) vectors of the InAs matrix. This indicates that this QD possesses a lattice mismatch strain (and elastic mismatch strain energy) minimizing orientation relationship with the matrix, as one would expect it to have if it originated from a structural transition. For this mixed anions III–V QD system, the results of the two already mentioned CTEM investigations[18,19] on the same TEM specimen that were separated by approximately 2 years in time also suggested that the number density of ordinarily strained QDs declines over time. Again, a simple explanation of these observations is structural transitions from ordinarily strained QDs to atomically ordered QDs over time.

To summarize the experimental observations, the only commonality of the samples reported upon in the original papers[21–28] was that a few years have elapsed between the growth and the TEM/STEM and PL analyses. The weight of the presented TEM/STEM observations (and the careful exclusion of possible specimen preparation and imaging artifacts[b]) in semiconductor QDs of different materials systems with either mixed cations or mixed anions, grown by different methods, leads this author to the working hypotheses that there are indeed morphological transformations and structural transitions in ordinarily strained random semiconductor alloy QDs over time.

[d]The high energy PL peaks in Fig. 5b are assigned as follows: 2.78 eV is due to a superposition of the (essentially) 2-D remains of the CdSe submonolayers and the excitonic transition in the ZnSe buffer and capping layers, 2.64 eV is due to a large population of small quasi 2-D CdSe-rich platelets (with sphalerite structure) that act as a first kind of ordinarily strained QD, and 2.25 eV is due to a small population of medium-sized 3-D CdSe-rich QDs (with sphalerite structure) that act as a second kind of ordinarily strained QD Ref. [28]. Such a coexistence of small quasi 2-D and medium-sized 3-D ordinarily stained QDs has actually been observed in similar samples of the same QD system Ref. [23]. The PL peak at 2.64 eV also shows the typical low-energy tail that is characteristic of a nonuniform Cd distribution in an epitaxial layer that contains small quasi 2-D CdSe-rich platelets Ref. [36].

So far, only a few in situ thermal treatments and structural analyses were performed in the electron microscopes under elevated temperatures, e.g., Fig. 2b. When elevated temperatures and annealing times parameter sets are found that in effect correspond to room temperature anneals over a few years, such experiments have the capacity to directly proof that there are structural transitions in ordinarily strained random alloy semiconductor QDs over time. On the other hand, such experiments are quite difficult to set up as atomically ordered QDs may exist only in certain approximate stoichiometries, which uncontrolled mismatch strain energy-driven out-diffusion of atoms from the ordinarily strained QDs during annealing may prevent from occurring. In addition, there must be a certain critical temperature above which randomness in the semiconductor alloy will prevail because of its larger entropy.

Nevertheless, the preliminary results are encouraging and suggest that the atomically ordered QD structure in Fig. 2b is thermodynamically more stable than the random alloy sphalerite-type structure in which this QD probably had grown a few years ago (Fig. 1a). More such experiments are to be performed in the future because it is expected that this will lead to both irrefutable evidence for the occurrence of the structural transitions that are currently only inferred from TEM/STEM observations [22–32] and to information on the operating mechanisms of atomic rearrangements in ordinarily strained random semiconductor alloy QDs.

As one can easily estimate,[26,29] there is an essentially hydrostatic pressure in the 1–10 GPa range on epitaxial ordinarily strained random semiconductor alloy QD and an excess Gibbs free energy of approximately 0.1–1 eV (i.e., $\gg kT$ at room temperature) for every atom of such a QD associated with the typically encountered lattice mismatch strains.[e] In addition, grown-in and quenched-in vacancies that remained from the cooling down of the QD structure after the growth will, in strain fields, move to the spatial positions of highest compressive strain,[40] which are exactly the positions of the Stranski–Krastanow grown QDs. When these vacancies arrive at the position of an ordinarily strained random semiconductor alloy QD, they can facilitate atomic rearrangements there that lead to a reduction of excess Gibbs free energy which results form the compressive stress on these QDs. It is thus not at all inconceivable that thermodynamics-driven morphological transformations and structural transitions should be possible over a long-enough time, even at room temperature.

While it can be seen directly in the TEM and STEM images (Figs. 2b, 3b, 4a, 5a, and 6a) that the size condition i and the condition of absence of detrimental defects (such as dislocations, iii) are fulfilled, atomically ordered semiconductor alloys (i.e., crystallographic superlattices) and phase-separated semiconductor alloys (i.e., chemical superlattices) are known to possess smaller band gaps than the random semiconductor alloys of the same net chemical compositions.[41] Combinations of HRTEM and photoluminescence spectroscopy, e.g., Ref. [28] or scanning tunneling microscopy-induced luminescence[42,43] investigations demonstrated that smaller band gaps than the random semiconductor alloys of the same net chemical composition are realized for atomically ordered semiconductor entities.

Failure of optoelectronic devices on the timescale of a few years may be explained by structural transitions in ordinarily strained QDs as discussed in this part because devices which employ less severely strained random semiconductor alloy QDs have been found to possess longer lifetimes.[5] Growing ordinarily strained, but thermodynamically unstable, random semiconductor alloy QDs that are under compressive stresses in the GPa range, structurally transforming these entities by means of suitable postgrowth thermal treatments, and recovering these novel structures to ambient or device-operating conditions may, on the other hand, become a way to realize Feynman's 1959 vision[44] of creating and exploiting atomic arrangements on the nanometer scale that nature does not otherwise provide.

CONCLUSION

Epitaxial semiconductor quantum dots as grown in the Stranski–Krastanow mode are typically alloyed, possess a more or less random distribution of the cations and/or anions over their respective sublattices, have a spatially nonuniform chemical composition distribution, and are compressively strained to a few percent. The lattice mismatch strains are believed to trigger atomic rearrangements inside quantum dots when the respective samples are stored at room temperature over time periods of a few years. These atomic rearrangements result in long-range atomic order and/or phase separation. While the results suggest that Stranski–Krastanow grown semiconductor quantum dots of certain III–V and II–VI systems are structurally unstable and that devices based on these quantum dots may fail over time, triggering and controlling structural transitions in epitaxial semiconductor quantum dots may also offer an opportunity

[e]For hydrostatic pressure, the product of the bulk modulus (order of magnitude 100 GPa for semiconductors) and the relative elastic volume change is equal to the product of pressure and volume, which is also the excess Gibbs free energy due to lattice mismatch strains. Starting with an unstrained sphere of 10-nm diameter with a lattice constant of 0.5 nm that contains 33510 atoms, one obtains for an elastic volume change of 5% a pressure of 5 GPa and an excess Gibbs free energy of 0.49 eV per atom. Hydrostatic pressures in this range are known to lead to structural transitions in elemental (group IV) and binary III–V compound semiconductors Ref. [37]. The usage of this simple formula is justified by studies on the size dependency of elastic properties of nanometer-sized particles, e.g., Refs. [38,39].

of creating atomic arrangements that nature does not otherwise provide.

ACKNOWLEDGMENTS

Collaborations with N.D. Browning, T. Topuria (both from the University of Illinois at Chicago at that time), G.R. Booker, Robin J. Nicholas (both from the University of Oxford/U.K.), Nigel J. Mason (also from the University of Oxford/U.K. at that time), and K. Pierz (Physical-technical Federal Institute, "PTB," Braunschweig/Germany) are kindly acknowledged. Fig. 5b is courtesy of K. Pierz. J.K. Furdyna and M. Dobrowolska (both from the University of Notre Dame) are thanked for the supply of samples. This research was supported by an award from Research Corporation.

REFERENCES

1. Bimberg, D. Quantum dots: Paradigm changes in semiconductor physics. Semiconductors **1999**, *33*, 951–955.
2. Bimberg, D.; Grundmann, M.; Ledentsov, N.N. *Quantum Dot Heterostructures*; John Wiley & Sons: Chichester, 1999.
3. Pearsall, T.P. *Quantum Semiconductor Devices and Technologies*; Kluwer Academic Publishers: Boston, 2000.
4. Ledentsov, N.N.; Ustinov, V.M.; Shchukin, V.A.; Kop'ev, P.S.; Alferov, Zh.I.; Bimberg, D. Quantum dot heterostructures: Fabrication, properties, lasers (review). Semiconductors **1998**, *32*, 343–365.
5. Liu, H.-Y.; Xu, B.; Wei, Y.-Q.; Ding, D.; Han, Q.; Liang, J.-B.; Wang, Z.-G. High-power and long-lifetime InAs/GaAs quantum-dot laser at 1080 nm. Appl. Phys. Lett. **2001**, *79*, 2868–2870.
6. Towe, E.; Pan, D. Semiconductor quantum-dot nanostructures: Their application in a new class of infrared photodetectors. IEEE J. Sel. Top. Quantum Electron. **2000**, *6*, 408–421.
7. Snider, G.; Orlov, A.O.; Kummamuru, R.K.; Ramasubramaniam, R.; Amlani, I.; Bernstein, G.H.; Lent, C.S. Quantum-dot cellular automata. Mater. Res. Soc. Symp. Proc. **2002**, *696*, N7.6.1–N7.6.11.
8. Shumway, J.; Williamson, A.J.; Zunger, A.; Passaseo, A.; DeGiorgi, M.; Cingolani, R.; Catalano, M.; Crozier, P. Electronic structure consequences of In/Ga composition variations in self-assembled In_xGa_{1-x} As/GaAs alloy quantum dots. Phys. Rev., B **2001**, *64*, 125302-1–125302-11.
9. Mahajan, S.; Harsha, K.S. *Principles of Growth and Processing of Semiconductors*; WCB/McGraw-Hill: Boston, 1999.
10. Stranski, I.N.; Krastanow, L. Zur Theory der orientierten Ausscheidung von Ionenkristallen aufeinander. Sitz.ber. Kais. Akad. Wiss., Math.-Nat.wiss. Cl., Abt. IIb **1938**, *146*, 797–818.
11. Walther, T.; Cullis, A.G.; Norris, D.J.; Hopkinson, M. Nature of the Stranski–Krastanow transition during epitaxy of InGaAs on GaAs. Phys. Rev. Lett. **2001**, *86*, 2381–2384.
12. Cullis, A.G.; Norris, D.J.; Walther, T.; Migliorato, M.A.; Hopkinson, M. Stranski–Krastanow transition and epitaxial island growth. Phys. Rev., B **2002**, *66*, 81305–81308.
13. Cullis, A.G.; Norris, D.J.; Walther, T.; Migliorato, M.A.; Hopkinson, M. Epitaxial island growth and the Stranski–Krastanow transition. Mater. Res. Soc. Symp. Proc. **2002**, *696*, N1.1.1–N1.1.8.
14. Rosenauer, A.; Oberst, W.; Litvinov, D.; Gerthsen, D.; Förster, A.; Schmidt, R. Structural and chemical investigation of $In_{0.6}Ga_{0.4}As$ Stranski–Krastanow layers buried in GaAs by transmission electron microscopy. Phys. Rev., B **2000**, *61*, 8276–8288.
15. Kret, S.; Ruterana, P.; Rosenauer, A.; Gerthsen, D. Extracting quantitative information from high resolution electron microscopy. Phys. Status Solidi, B **2001**, *227*, 247–295.
16. García, J.M.; Medeiros-Ribeiro, G.; Schmidt, K.; Ngo, T.; Feng, J.L.; Lorke, A.; Kotthaus, J.; Petroff, P.M. Intermixing and shape changes during the formation of InAs self-assembled quantum dots. Appl. Phys. Lett. **1997**, *71*, 2014–2016.
17. Möck, P.; Booker, G.R.; Mason, N.J.; Alphandéry, E.; Nicholas, R.J. MOVPE grown self-assembled Sb-based quantum dot assessed by means of AFM and TEM. IEE Proc., Optoelectron. **2000**, *147*, 209–215.
18. Möck, P.; Booker, G.R.; Alphandéry, E.; Nicholas, R.J.; Mason, N.J. Self-assembled InSb quantum dots in InAs and GaSb matrices assessed by means of TEM, AFM and PL. Inst. Phys. Conf. Ser. **1999**, *164*, 133–136.
19. Norman, A.G.; Mason, N.J.; Fisher, M.J.; Richardson, J.; Krier, A.; Walker, P.J.; Booker, G.R. Structural and optical characterization of MOVPE self-assembled InSb quantum dots in InAs and GaSb matrices. Inst. Phys. Conf. Ser. **1997**, *157*, 353–356.
20. Kim, C.S.; Kim, M.; Furdyna, J.K.; Dobrowolska, M.; Lee, S.; Rho, H.; Smith, L.M.; Jackson, H.E.; James, E.M.; Xin, Y.; Browning, N.D. Evidence for 2-D precursors and interdiffusion in the evolution of self-assembled CdSe quantum dots on ZnSe. Phys. Rev. Lett. **2000**, *85*, 1124–1127.
21. Möck, P.; Booker, G.R.; Mason, N.J.; Nicholas, R.J.; Alphandéry, E.; Topuria, T.; Browning, N.D. MOVPE grown self-assembled and self-ordered

InSb quantum dots in a GaSb matrix assessed by AFM, CTEM, HRTEM and PL. Mater. Sci. Eng., B **2001**, *80*, 112–115.

22. Möck, P.; Topuria, T.; Browning, N.D.; Dobrowolska, M.; Lee, S.; Furdyna, J.K.; Booker, G.R.; Mason, N.J.; Nicholas, R.J. Internal self-ordering in In(Sb,As), (In,Ga)Sb, and (Cd,Zn,Mn)Se nano-agglomerates/quantum dots. Appl. Phys. Lett. **2001**, *79*, 946–948.
23. Möck, P.; Topuria, T.; Browning, N.D.; Titova, L.; Dobrowolska, M.; Lee, S.; Furdyna, J.K. Self-ordered CdSe quantum dots in ZnSe and (Zn,Mn)Se matrices assessed by transmission electron microscopy and photoluminescence spectroscopy. J. Electron. Mater. **2001**, *30*, 748–755.
24. Möck, P.; Topuria, T.; Browning, N.D.; Booker, G.R.; Mason, N.J.; Nicholas, R.J.; Titova, L.V.; Dobrowolska, M.; Lee, S.; Furdyna, J.K. Self-ordering in CdSe/ZnSe, CdSe/(Zn,Mn)Se, InSb/GaSb and InSb/InAs quantum dot structures and a novel type of quantum dot. Mater. Res. Soc. Symp. Proc. **2001**, *642*, J6.3.1–J6.3.6.
25. Topuria, T.; Möck, P.; Browning, N.D.; Titova, L.V.; Dobrowolska, M.; Lee, S.; Furdyna, J.K. Z-contrast transmission electron microscopy on self-assembled CdSe quantum dots in ZnSe and (Zn,Mn)Se matrices. Mater. Res. Soc. Symp. Proc. **2001**, *642*, J8.3.1–J8.3.6.
26. Möck, P.; Topuria, T.; Browning, N.D.; Nicholas, R.J.; Booker, G.R. Atomic self-ordering in heteroepitaxially grown semiconductor quantum dots due to relaxation of external lattice mismatch strains. Mater. Res. Soc. Symp. Proc. **2002**, *696*, N8.8.1–N8.8.6.
27. Möck, P.; Lei, Y.; Topuria, T.; Browning, N.D.; Ragan, R.; Min, K.S.; Atwater, H.A. Structural Transformations in Self-Assembled Semiconductor Quantum Dots as Inferred by Transmission Electron Microscopy. In *Physical Chemistry of Interfaces and Nanomaterials*, Proceedings of the Society of Photo-Optical Instrumentation Engineers, Seattle, WA, July 7–9, 2002; Vol. 4807.
28. Möck, P.; Pierz, K.; Topuria, T.; Browning, N.D.; Wu, H.; McCann, P. Atomic ordering in self-assembled epitaxial II–VI and IV–VI compound semiconductor quantum dot systems. Mater. Res. Soc. Symp. Proc. **2003**, *749*, W13.5.1–W13.5.6.
29. Möck, P. Atomic ordering in self-assembled epitaxial and endotaxial compound and element semiconductor quantum dot structures: The first review. Mater. Res. Soc. Symp. Proc. **2003**, *776*, Q5.4.1–Q5.4.6.
30. Browning, N.D.; Arslan, I.; Moeck, P.; Topuria, T. Atomic resolution scanning transmission electron microscopy. Phys. Status Solidi, B **2001**, *227* (1), 229–245.
31. Browning, N.D.; Arslan, I.; James, E.M.; Moeck, P.; Topuria, T.; Xin, Y. Analyzing interfaces and defects in semiconducting materials on the atomic scale. Inst. Phys. Conf. Ser. **2001**, *169*, 1–12.
32. Browning, N.D.; Arslan, I.; Ito, Y.; James, E.M.; Klie, R.; Moeck, P.; Topuria, T.; Xin, Y. Application of atomic scale STEM techniques to the study of interfaces and defects in materials. J. Electron Microsc. **2001**, *50* (3), 205–218.
33. Srivastava, G.P.; Martins, J.L.; Zunger, A. Atomic structure and ordering in semiconductor alloys. Phys. Rev. **1985**, *B31*, 2561–2564. Erratum Phys. Rev. **1988**, *B38*, 12694.
34. O'Neill, J.P.; Ross, I.M.; Cullis, A.G.; Wang, T.; Parbrook, P.J. Electron-beam-induced segregation in InGaN/GaN multiple-quantum wells. Appl. Phys. Lett. **2003**, *83*, 1965–1967.
35. Johnson, W.C. Influence of Elastic Stress on Phase Transformations. In *Lectures on the Theory of Phase Transformations,* 2nd Ed.; Aaronson, H.I., Ed.; The Minerals, Metals & Materials Society: Warrendale, 2001; 35–134.
36. Kim, M.; Furdyna, J.K.; Dobrowolska, M.; Lee, S.; Cheon, M.; Luo, H. Transition form island formation to pseudomorphic growth in the submonolayer CdSe/ZnSe multilayer system. Appl. Phys. Lett. **2003**, *83*, 1728–1730.
37. Ackland, G.J. High-pressure phases of group IV and III–V semiconductors. Rep. Prog. Phys. **2001**, *64*, 483–516.
38. Baker, S.P.; Vinci, R.P.; Arias, T. Elastic and anelastic behavior of materials in small dimensions. MRS Bull. **2002**, *27* (1), 26–29.
39. Miller, R.E.; Shenoy, V.B. Size-dependent elastic properties of nanosized structural elements. Nanotechnology **2000**, *11*, 139–147.
40. Girifalco, L.A.; Welch, D.O. *Point Defects and Diffusion in Strained Metals*; Gordon and Breach: New York, 1967.
41. Zunger, A.; Mahajan, S. Atomic Ordering and Phase Separation in Epitaxial III–V Alloys. In *Handbook on Semiconductors,* Completely Revised Ed.; Moss, T.S., Mahajan, S., Eds.; Elsevier Science B.V.: Amsterdam, 1994; Vol. 3, 1399–1514.
42. Håkanson, U.; Sass, T.; Johansson, K.-M.; Pistol, M.-E.; Samuelson, L. Quantum-dot-induced ordering in $Ga_xIn_{1-x}P$/InP islands. Phys. Rev., B **2002**, *66*, 235308-1–235308-5.
43. Håkanson, U.; Zwiller, V.; Johansson, K.-M.; Sass, T.; Samuelson, L. Luminescence polarization of ordered GaInP/InP islands. Appl. Phys. Lett. **2003**, *82*, 627–629.
44. Feynman, R.P. In *There is Plenty of Room at the Bottom*, Lecture at the Annual Meeting of the American Physical Society, December 29, 1959. http://www.zyvex.com/nanotech/feynman.html.

Quantum Dots, Semiconductor: Site-Controlled Self-Organization

S. Kohmoto
H. Nakamura
S. Nishikawa
T. Yang
K. Asakawa
The Femtosecond Technology Research Association (FESTA), Ibaraki, Japan

INTRODUCTION

The spontaneous formation of nanoscale islands in strained-layer epitaxy, represented by Stranski–Krastanov mode growth, has been widely studied as a simple and efficient method for producing semiconductor quantum dots.[1–3] In such a self-organizing approach, quantum dots form randomly on surfaces. If the formation sites of individual quantum dots can be specified, they can be arranged in various configurations, from a single dot to regular, dense, and even coupled dot arrays, and dot placement can be restricted to specific regions, enabling the development of novel electronic and optical devices.

This article proposes a technique for controlling the self-organization sites of individual quantum dots using scanning tunneling microscope (STM) probe-assisted nanolithography and self-organizing molecular beam epitaxy (MBE).[4–10] Two-dimensional (2-D) and three-dimensional (3-D) arrays of InAs site-controlled quantum dots (SCQDs) are successfully fabricated on GaAs substrates with nanoscale pitch and precision. Photoluminescence from the SCQD arrays is investigated at room temperature. The tip repositioning function of the STM system reveals the growth mode of the SCQDs.

EXPERIMENT

SCQD fabrication was done in situ in an ultrahigh vacuum multichamber system[6,7,9] to achieve high-quality structures. As illustrated in Fig. 1, this system consists of chambers for STM (Large Sample STM, Omicron), solid-source MBE, Auger electron spectroscopy (AES), STM tip cleaning, surface processing, and sample loading, which are connected to each other via a 3-m-long ultrahigh vacuum transfer tunnel. The tungsten (W) probes used for both STM-assisted SCQD fabrication and STM measurements were prepared by electrochemical etching in NaOH solution and cleaned by electron beam heating in the tip-cleaning chamber.

The STM apparatus is specially designed to have a unique tip repositioning function.[6,7] This function enables the step-by-step observation of the formation of individual quantum dots. For this purpose, arrays of 30×30 μm^2 mesa were patterned on n-GaAs(001) substrates by conventional photolithography and chemical wet etching. After growth of GaAs buffer layers, quantum dot arrays having in-plane array sizes of 1×1–2×2 μm^2 were fabricated at the centers of the mesa top surfaces with a positional accuracy of better than several micrometers. This accuracy was realized by monitoring the approach of the STM probe tip to the mesa top surfaces using a long working distance, high-resolution optical microscope (QM100, Questar; spatial resolution of about 2 μm at a distance of 20 cm) from outside the STM chamber. These dot arrays were easily and repeatedly found in later STM measurements, even when the sample was transferred between the STM and MBE chambers, because of the large scanning area (up to 15×15 μm^2) of our STM, which enabled detection of the target dot array in a single scan (the typical scan area is 4×4 μm^2). This technique thus permitted the monitoring of the evolution of an identical quantum dot in the target dot array. STM images were taken at sample bias voltages of -2.4 to -5.1 V and tunneling currents of 0.05–0.49 nA with constant-current feedback. The optical properties of the quantum dots were investigated by microprobe photoluminescence at room temperature. Photocarriers were generated by He–Ne laser (633 nm) of 3-μm beam diameter, and the photoluminescence signal was detected with a cooled InGaAs photodiode array.

RESULTS AND DISCUSSION

Nanoscale Site Control

Fig. 2 schematically illustrates the SCQD fabrication procedure. This procedure comprises in situ deposition and growth, without complex processes such as mask layer formation and pattern etching. First, a flat and clean

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120015485

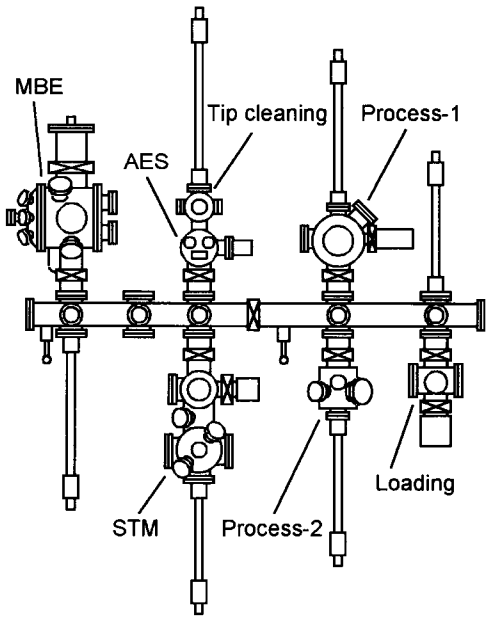


Fig. 1 Schematic illustration of the ultrahigh vacuum–STM/MBE multichamber system. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

GaAs(001) surface was prepared by MBE, and the STM W tip was located at a surface target position with a sample bias voltage of -3.2 V and a tunneling current of 0.1 nA. Then, (a) nanoscale deposits were created on a GaAs surface by applying 3–5 pulses of voltage and current (+7 to +8 V and 10 nA for 500 msec) at intervals of 10 msec between the surface and the W probe. The primary composition of the deposit is considered to be a W or a W-containing compound, because the STM probe is composed of W. We confirmed that the deposits remain stable at temperatures up to at least 610°C under arsenic pressure and act as "nanomasks" on which GaAs does not grow directly. Accordingly, when a thin layer of GaAs or GaAs/Al(Ga)As superlattice was subsequently grown on this surface, (b) the GaAs or superlattice tended to avoid the nanomasks at the initial growth stage, but (c) later covered the nanomasks by lateral growth, leading to the natural formation of nanoholes just above the nanomasks. Finally, (d) the supply of InAs on this surface resulted in self-organization of SCQDs at the nanohole sites.

Fig. 3a, b, and c shows STM images for the fabrication process of an SCQD array corresponding to the steps in Fig. 2a, c, and d, respectively. These are step-by-step images of the identical surface region obtained using the tip repositioning function. As shown in Fig. 3a, a 2×4 nanomask array was initially created on the GaAs surface. All nanomasks had a similar size of 3 nm in height and 16 nm in base diameter. Fig. 3b shows GaAs nanoholes produced by subsequent 15-nm-thick GaAs growth at 460°C. The holes were elongated in the [110] direction, because of different incorporation rates of Ga atoms in different surface planes, and had an area of 27×42 nm^2 and a depth of 4 nm. Comparisons of nanomask height, GaAs layer thickness, and hole depth indicated that the namomasks were completely covered with the GaAs. Then, as shown in Fig. 3c, a 2×4 quantum dot array was self-organized exactly at the nanohole sites by 1.1-ML InAs supply at 460°C. In this dot growth, a 0.17 monolayer [ML; 1 ML corresponds to the surface atom density of GaAs(001)] supply per 4 sec (occasionally, 0.087 ML per 2 sec) was repeated with a growth interruption of 1 min under a continuous arsenic flux of 1.3×10^{-5} Torr. We observed virtually no undesirable Stranski–Krastanov quantum dots (SKQDs), the occurrence of which became obvious at 1.3 ML, in the flat surface region. This indicates highly selective SCQD formation. The SCQD was 6 nm in height and about 35 nm in base diameter. A magnified STM image of the SCQD revealed a hexagonal base slightly elongated in the [-110] direction and a faceted dot surface.

STM-induced nanomasks can be created by this method with nanoscale pitch and precision, and the GaAs nanoholes defined above nanomasks are similar in size to SCQDs. Therefore, the resulting SCQDs can be located in close proximity comparable to the dot diameter. Fig. 4 demonstrates such close dot positioning, showing three SCQD pairs fabricated by the present site-control method. A 100-nm pitch array of paired SCQDs with 45-nm center-to-center distance and 15-nm bottom-edge spacing was successfully produced.

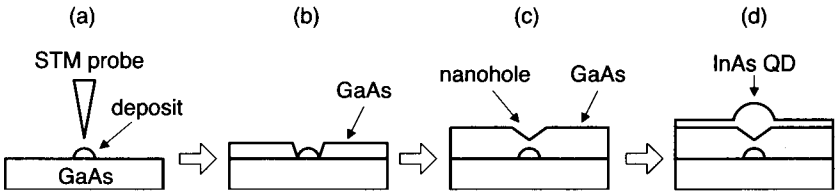


Fig. 2 Fabrication procedure for InAs SCQDs on GaAs(001) surfaces using STM probe-assisted nanolithography and self-organizing MBE. (a) STM-induced deposits (nanomasks), (b) and (c) GaAs nanoholes, and (d) InAs SCQDs. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

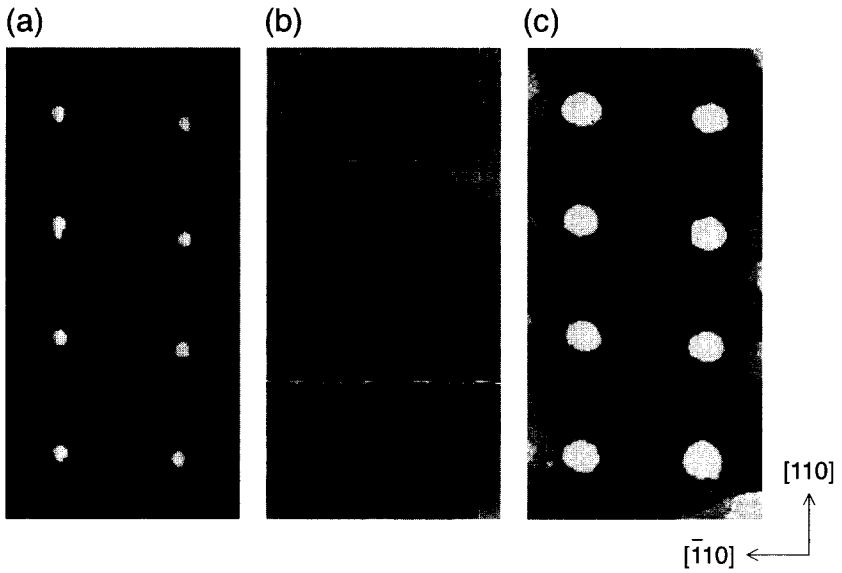


Fig. 3 Step-by-step STM images for the fabrication process of a 2×4 InAs SCQD array on GaAs surfaces corresponding to Fig. 2a, c, and d. (a) STM-induced deposits (nanomasks; height 3 nm, base diameter 16 nm), (b) GaAs nanoholes (depth 4 nm, size 27×42 nm), and (c) InAs SCQDs (height 6 nm, base diameter 35 nm). These are images of the identical surface region obtained using the tip repositioning function. Image area: 200×400 nm^2. (Reprinted with permission from Ref. [4]. Copyright 1999, American Institute of Physics.)

Mechanism of Site-Controlled Self-organization

In previous studies, we used submicron-sized GaAs holes to site-control InAs dots.[11,12] In those cases, the dots selectively appeared in the holes at the critical InAs amount for 3-D islanding in Stranski–Krastanov growth mode. The selectivity arose from preferential InAs nucleation at specific planes in the holes, which have a high In incorporation rate. However, in the current study, the SCQDs at the nanohole sites were found to occur in a unique manner, unlike the above case, as described below.

Fig. 5 shows other detailed STM images of the SCQD formation process.[8] These are images of the identical surface region obtained using the tip repositioning function. Fig. 5a shows nanoholes produced by growing a 21-nm-thick GaAs/AlGaAs superlattice layer (1.5-nm GaAs/1.5-nm $Al_{0.3}Ga_{0.7}As \times 6$, capped with 3-nm GaAs) at 620°C on the surface with the STM probe-induced nanomasks. The holes have an area of 40×55 nm^2 and a depth of 10 nm. Also seen in the flat region between the holes are terraces and 2-D islands of a single step height (0.28 nm) elongated in the [−110] direction. This anisotropic shape occurs because the incorporation probability of Ga and Al atoms at B steps (parallel to [110]) is greater than that at A steps (parallel to [−110]) during superlattice growth.[13] Fig. 5b shows an STM image of the identical surface region after 0.7 ML InAs is supplied. The nanoholes are filled and almost planarized, but the filled parts exhibit slightly bright contrast in the STM image, as indicated, for instance, by arrowhead "a" in Fig. 5b. The area of contrast is elliptical in accordance with the shape of the original holes, and the height corresponding to the contrast is less than that of a single step. Most of the terraces and 2-D islands increased in size, although some of the relatively small 2-D islands disappeared. Consequently, the enlarged terraces and 2-D islands on the same level

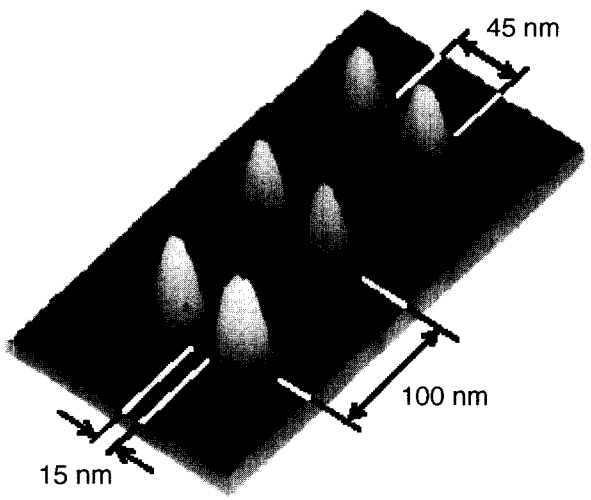


Fig. 4 STM image of three InAs SCQD pairs on GaAs. Each quantum dot is 6 nm in height and 30 nm in base diameter. Image area: 155×300 nm^2. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

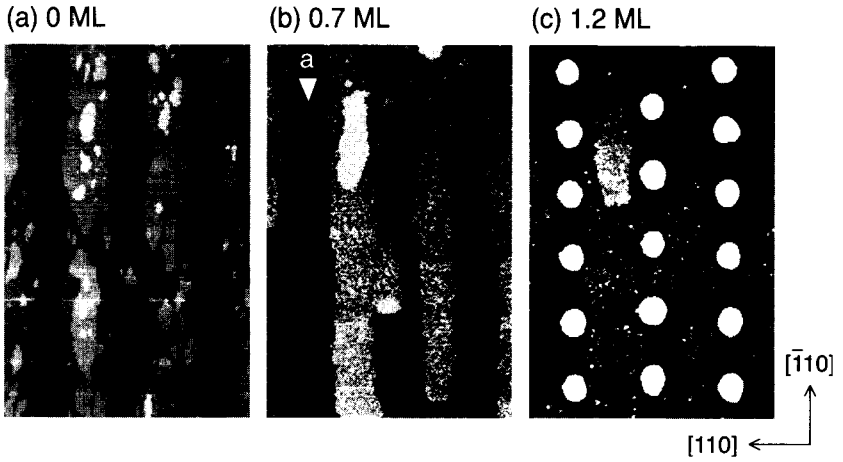


Fig. 5 STM images showing SCQD and wetting layer evolution with InAs supply. InAs: (a) 0 ML, (b) 0.7 ML, and (c) 1.2 ML. The white arrowhead "a" indicates the position of the original hole. These are images of the identical surface region obtained using the tip repositioning function. Image area: 380×580 nm^2. (Reprinted from Ref. [8]. Copyright 2002, with permission from Elsevier Science.)

coalesced into one larger terrace or 2-D island elongated in the $[-110]$ direction. Some terraces at the lower level disappeared as they were covered by the upper terraces. It should be noted that the extending terraces covered most of the filled hole sites in this process. At 1.2-ML InAs, as shown in Fig. 5c, quantum dots were selectively self-organized at the filled hole sites. In the flat region, terrace size further increased, with the result that only three levels of terraces dominated the surface.

In Fig. 5, we assume that the material filling the nanoholes is InAs or InGaAs and that the observed extension of terraces and 2-D islands indicates the formation of an In(Ga)As wetting layer (WL) that covers the filled hole sites prior to SCQD formation. In this case, since the filling In(Ga)As has a larger lattice constant and lower bandgap energy than the surrounding GaAs, the slightly bright contrast at the filled hole sites in the constant-current STM image shown in Fig. 5b can be explained by upheaval due to compression and/or larger tunneling conductivity at the wetting layer surfaces on the filled holes. Then, the subsequently supplied In atoms for 3-D islanding accumulate at the filled hole sites, because the lattice mismatch of InAs with the wetting layer on the filling In(Ga)As is less than that on the surrounding GaAs, leading to selective self-organization of quantum dots. This growth process can be understood by analogy with vertically aligned SKQDs in close stacking.[14] A detailed analysis of the evolution of the SCQD and wetting layer is provided in another publication.[5]

In order to form the quantum dots at the nanohole sites, it is important to both cover the nanomasks completely with GaAs (or superlattices) and ensure that the holes are deeper than a certain critical value by selecting suitable GaAs growth conditions and thicknesses. For instance, although 5-nm-thick GaAs growth created 25×35-nm^2-sized holes, the 1.2-ML InAs supply on this surface did not result in dot formation at the nanohole sites. This is because the top part of the nanomask, on which InAs does not grow directly, was still exposed at the hole bottom because of the thin GaAs layer. On the other hand, even when the nanomasks under certain growth conditions were completely covered with GaAs, dots did not form at nanohole sites when the hole depth was less than about 3 nm. In this case, because the volume of In(Ga)As filling the holes was small, the lattice constant of the wetting layer surface on the filling In(Ga)As was presumably so close to that on the surrounding GaAs surface that In atoms could not selectively accumulate at the nanohole sites.

Three-Dimensional Site Control

When SKQD layers are stacked with thin spacer layers, the dots align vertically[14,15] or obliquely,[16,17] depending on elastic anisotropy and the thickness of the spacers, because of strain propagation. It has been predicted that regular strain distribution will gradually develop on the spacer-layer surfaces with increasing stack numbers, leading to increasingly uniform size for the topmost dots.[14,16–19] If the stacking is started with regular SCQD arrays in place of random SKQDs, the dots can be site controlled into 3-D quantum dot lattices, and uniform SKQDs array will be obtained more efficiently. In addition, this 3-D site control has the potential to realize novel quantum dot structures such as 3-D quantum dot crystals[16] with perfect spatial ordering. In this section,

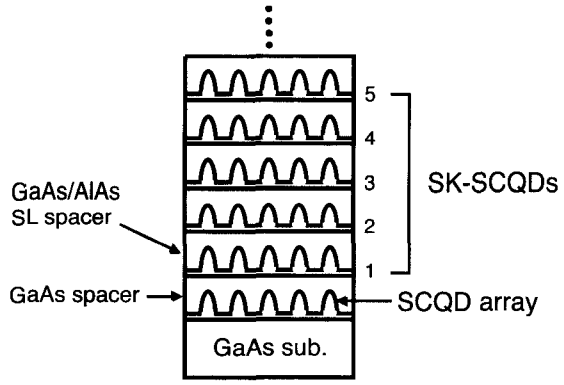


Fig. 6 Schematic illustration of a 3-D InAs SK-SCQD lattice structure. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

we demonstrate 3-D site control of SKQDs based on the site-initiated multistacking approach[9] mentioned above.

Fig. 6 illustrates the structure of a 3-D InAs quantum dot lattice. In this structure, an InAs SCQD array of arbitrary configuration was initially prepared on GaAs surfaces by the STM probe-assisted site-control technique described earlier. This array defines the in-plane lattice symmetry and parameters of the 3-D dot lattice. Then, after growth of a 10-nm-thick GaAs spacer layer, InAs was supplied to form SKQDs just above the arrayed SCQDs by strain-induced preferential self-organization. Hereafter, we denote the SKQDs grown above the SCQDs as Stranski–Krastanov mode-grown site-controlled quantum dots (SK-SCQDs). Repeating the spacer layer [GaAs/AlAs superlattice (SL) in case of Fig. 6] and SK-SCQD growth permitted the formation of a 3-D quantum dot lattice with vertically and horizontally aligned dots.

Fig. 7 shows STM images for the vertical stacking process of the SK-SCQDs in the 3-D quantum dot lattice.[9] These are images of the identical surface region. The first-layer SK-SCQDs shown in Fig. 7a were grown with 0.9 ML InAs supply. This amount is 0.4 ML less than the critical value for the formation of conventional InAs SKQD on unstrained GaAs surfaces, indicating the existing surface strain field on the spacer layer grown over the SCQDs. The SK-SCQD array exhibits a square lattice, although the alignment is not perfect and some dots are missing due to the initial lack of SCQDs. The lattice has unit vectors in the [100] and [010] directions and a lattice parameter of about 100 nm, which are defined by the embedded SCQD strain template. Fig. 7b shows the surface topography of a 20-nm-thick GaAs/AlAs superlattice spacer, terminated with GaAs, grown on the first SK-SCQD layer. The similar positions of surface steps in this image to those in Fig. 7a confirm that these two images are of the identical surface area. The spacer layer surface is almost planarized and dominated by monolayer-high terraces. However, the strain field at this surface, generated by the embedded SK-SCQDs, was visualized when InAs was grown in the next step. Fig. 7c shows an STM image of the second layer SK-SCQDs formed with 1.3 ML InAs supply on the spacer layer. As confirmed by comparing the positions of missing dots between Fig. 7a and c, the square lattice configuration of the first-layer dots is retained in the second-layer dots. The vertical pairing probability of dots between two layers is almost 100%. On the other hand, no dot is observed at interstitial positions. These results indicate that the spatial strain field can be engineered by the site-control technique.

The in-plane lattice symmetry and parameter of the 3-D quantum dot lattice can be arranged arbitrarily, independent of the crystallographic symmetries of the substrates, by

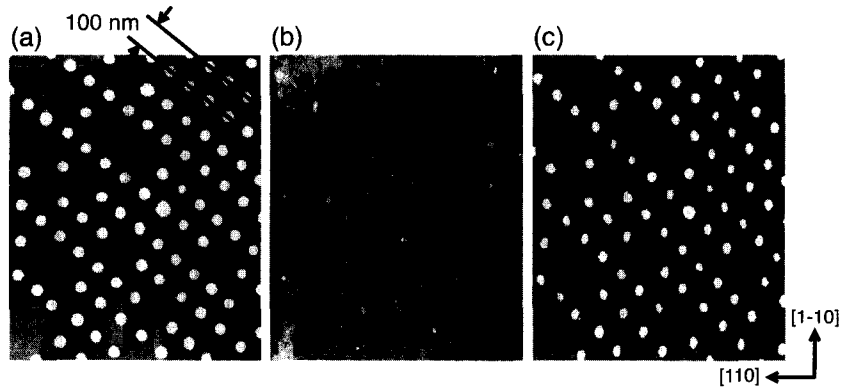


Fig. 7 STM images for vertical stacking process of InAs SK-SCQDs. (a) the first SK-SCQD layer, (b) spacer layer, and (c) the second SK-SCQD layer. These are images of the identical surface region obtained using the tip repositioning function. Image area: 1000×1200 nm^2. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

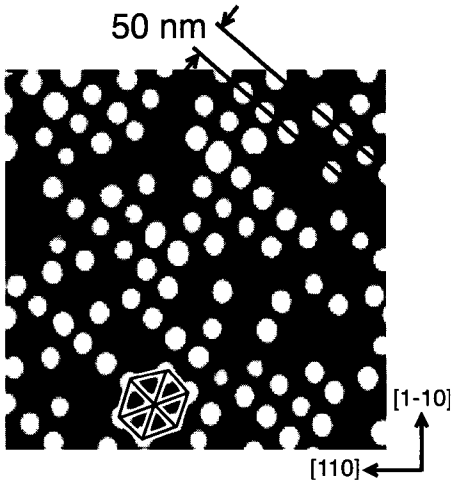


Fig. 8 STM image of a high-density InAs SK-SCQD array. Hexagonal-like quantum dot configuration is indicated. Image area: 600×600 nm^2. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

designing the SCQD strain template. An example is presented in Fig. 8, where the SK-SCQDs array has a hexagonal-like unit cell with a lattice parameter of 50–60 nm.[9] This configuration can achieve area dot densities as high as approximately 4×10^{10} cm^{-2}. Close positioning comparable to the dot diameter is realized through the nanoscale precision of the STM-assisted site-control method.

Optical Properties

Fig. 9a shows the microprobe-photoluminescence spectrum of a 3-D lattice of InAs SK-SCQDs measured at room temperature.[9] The sample comprised five SK-SCQD layers stacked in a manner similar to the structures in Fig. 7. The in-plane lattice parameter of the dot layer was about 100 nm. The total number of SK-SCQDs within the excitation laser beam was estimated to be about 1200 by STM observation of the top SK-SCQD layer before capping layer growth. As shown for the case "with SCQD template" in Fig. 9a, distinct photoluminescence from the 3-D SK-SCQD lattice is observed from 1.05 to 1.30 eV. The different photoluminescence experiments confirmed that the SCQDs of the strain templates exhibit no emission and do not contribute to the photoluminescence spectra of quantum dots in Fig. 9, probably because of the existence of the STM-induced nanomasks at close positions. When a reference area without the initial SCQD strain template on the same sample is excited, no photoluminescence from quantum dots is detected, as shown for the case "without SCQD template" in Fig. 9a. This means that conventional SKQDs are hardly formed in the stacked structure without the SCQD strain template, showing highly selective formation of the 3-D SK-SCQD lattices. The excitation power dependence of the photoluminescence spectral width indicated that the observed quantum dot emission is mainly due to ground state transitions in the SK-SCQDs. Photoluminescence intensity reduction for the WL and the GaAs barrier in the 3-D SK-SCQD lattice suggests that photocarriers were effectively consumed at dots through radiative recombination. Fig. 9b shows the photoluminescence spectrum of a two-layer stacked SK-SCQD sample fabricated in a manner similar to the five-layer sample. Although the number of dots per single layer was increased to some extent in this sample. the total number of measured dots fell to about 600. However, photoluminescence from the SK-SCQDs remains observable at room temperatures, as shown for the case "with SCQD template" in Fig. 9b. These photoluminescence results reveal the good crystallographic quality of the SK-SCQD structures, despite the addition of an artificial STM process and STM observation in some fabrication steps. This is because the present 3-D site control is carried out

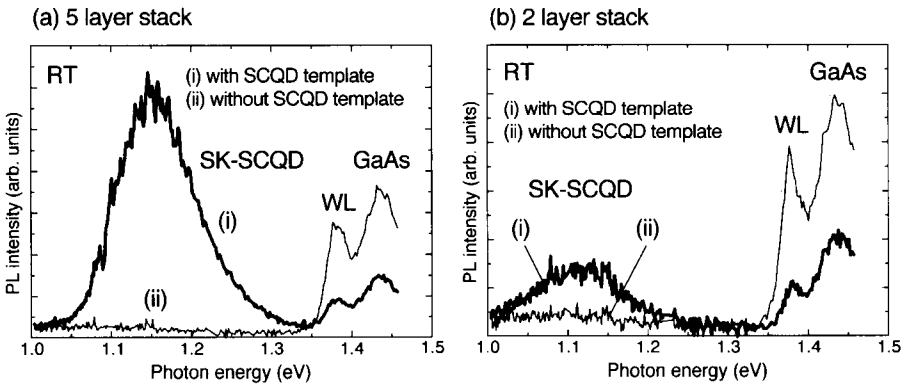


Fig. 9 Room temperature photoluminescence (PL) spectra of 3-D InAs SK-SCQD lattices. Samples include (a) five-layer stacked and (b) two-layer stacked SK-SCQDs. (Reprinted with permission from Ref. [9]. Copyright 2002, AVS—The Science & Technology Society.)

by ultrahigh vacuum in situ processing, which can keep the sample surface clean during the fabrication process,[20] using only the crystal growth technique after the initial nanomask deposition.

In Fig. 9, full width at half maximum of the SK-SCQD emissions is more than 100 meV, even in the two-layer stacked structure. This large line width can be attributed to differences in dot size in intradot layers and in interdot layers. For device applications, it is desirable to decrease both kinds of size differences to obtain a narrow line width. Targeting this improvement, we studied the photoluminescence of a single layer of SK-SCQDs.[10] A single layer of an InAs SK-SCQD square lattice with a lattice parameter of 75 nm was fabricated above an InAs SCQD strain template. The SK-SCQDs were grown at lower growth rate and lower arsenic pressure and with a greater supply of InAs than applied to the SCQDs during growth. This resulted in the formation of relatively large SK-SCQDs with an average height of about 9 nm and base diameter of about 45 nm. The SK-SCQDs were then buried with $GaAs/Al_{0.3}Ga_{0.7}As/GaAs$. Fig. 10 shows a photoluminescence spectrum from the buried structure at room temperature. From excitation spectral measurements, two respective low energy peaks were identified to arise from the ground and first excited states of the SK-SCQDs. The other two peaks, at about 1.36 and 1.44 eV, are, respectively, attributable to the WL and GaAs substrate. The full width at half maximum of the ground state peak for the SK-SCQDs is about 40 meV, which is fairly good for artificially arranged quantum dot arrays. This is because the SK-SCQDs, like conventional SKQDs,

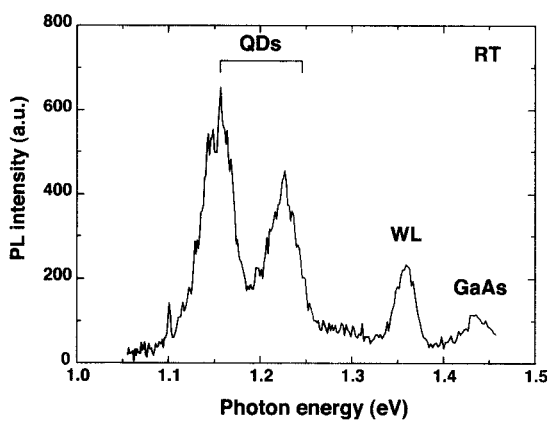


Fig. 10 Room temperature photoluminescence (PL) spectrum of a single-layer InAs SK-SCQD lattice with a pitch of 75 nm. The two peaks at lower photon energies are from the ground and first excited states of the InAs SK-SCQDs, respectively. The peak at about 1.36 eV is from the InAs wetting layer (WL). The peak at around 1.44 eV is from the GaAs substrate. (Reprinted with permission from Ref. [10]. Copyright 2003, American Institute of Physics.)

can be made higher under the specific growth conditions described above, and because the energy variation of such higher dots is relatively insensitive to size fluctuations.

CONCLUSION

We developed a site-controlled self-organization technique for InAs quantum dots on GaAs(001) surfaces using in situ STM probe-assisted nanolithography and self-organizing MBE. We found that it was possible to create nanoscale deposits on a GaAs surface by applying voltage and current pulses between the surface and an STM probe. When the deposits were used as nanomasks for GaAs growth, GaAs nanoholes were formed naturally above the deposits. Subsequently supplied InAs filled and planarized the GaAs nanoholes, providing flat surfaces with modulated strain fields. Consequently, supplying additional InAs on the surfaces led to SCQD formation at the filled-hole sites. Virtually no undesirable SKQDs were observed, indicating highly selective dot formation. This process resulted in successful fabrication of 2-D arrays of InAs SCQDs with nanoscale pitch and precision. 3-D quantum dot lattices were also demonstrated. In-plane dot arrangement of the 3-D dot lattices was initially defined by 2-D arrays of InAs SCQDs. With the SCQD arrays used as strain templates, self-organized InAs quantum dots were vertically aligned by strain-induced multistacking, resulting in 3-D quantum dot lattices. Photoluminescence measurements revealed good crystallographic quality for the 2-D and 3-D SK-SCQD lattices.

The site-control technique demonstrated here enables the location of an isolated single quantum dot at an intended position on the surfaces, which is useful for developing single-photon and electron devices. The high-density dot arrays with small energy dispersion are also technologically significant structures that can be achieved by this site-control approach, because large dots, whose energy variations are relatively insensitive to size fluctuations, can be formed regardless of density. This independent control of dot size and density is a difficult task for conventional crystal growth. Another attractive application of the site-control technique is the formation of dot arrays in limited areas in the devices. For instance, photonic crystal waveguide structures having quantum dot arrays at specific regions would make it possible to build miniaturized, ultrahigh speed all-optical switches with high switching efficiency.[21,22] Such applications will require improvements in the uniformity of the size of SK-SCQDs in the intra- and interlayers. In 3-D site-control processes, it is also important to control lateral strain interaction between neighboring dot sites in vertical stacking, as we pointed out in a recent study.[10] Further experimental investigations and theoretical consideration of growth conditions

and spatial strain fields will lead to uniform quantum dot size and the realization of novel nanostructure systems.

ACKNOWLEDGMENTS

This work was supported by the NEDO within the framework of the Femtosecond Technology Project.

REFERENCES

1. Goldstein, L.; Glas, F.; Marzin, J.Y.; Charasse, M.N.; Le Roux, G. Growth by molecular beam epitaxy and characterization of InAs/GaAs strained-layer superlattices. Appl. Phys. Lett. **1985**, *47* (10), 1099–1101.
2. Leonard, D.; Krishnamurthy, M.; Reaves, C.M.; Denbaars, S.P.; Petroff, P.M. Direct formation of quantum-sized dots from uniform coherent islands of InGaAs on GaAs surfaces. Appl. Phys. Lett. **1993**, *63* (23), 3203–3205.
3. Bimberg, D.; Grundmann, M.; Ledentsov, N.N. *Quantum Dot Heterostructures*; John Wiley & Sons: New York, 1999, and references therein.
4. Kohmoto, S.; Nakamura, H.; Ishikawa, T.; Asakawa, K. Site-controlled self-organization of individual InAs quantum dots by scanning tunneling probe-assisted nanolithography. Appl. Phys. Lett. **1999**, *75* (22), 3488–3490.
5. Nakamura, H.; Kohmoto, S.; Ishikawa, T.; Asakawa, K. Novel nano-scale site-controlled InAs quantum dot assisted by scanning tunneling microscope probe. Physica, E **2000**, *7* (3–4), 331–336.
6. Kohmoto, S.; Nakamura, H.; Ishikawa, T.; Asakawa, K. STM probe-assisted site-control of self-organized InAs quantum dots on GaAs surfaces. J. Electron. Mater. **2000**, *29* (5), 525–529.
7. Kohmoto, S.; Nakamura, H.; Ishikawa, T.; Asakawa, K. STM probe-assisted site-control of self-organized InAs quantum dots on GaAs surfaces. J. Electron. Mater. **2000**, *29* (9), 1110.
8. Kohmoto, S.; Nakamura, H.; Ishikawa, T.; Nishikawa, S.; Nishimura, T.; Asakawa, K. Site-controlled self-organization of InAs quantum dots. Mater. Sci. Eng., B **2002**, *88* (2–3), 292–297.
9. Kohmoto, S.; Nakamura, H.; Nishikawa, S.; Asakawa, K. Three-dimensional site control of self-organized InAs quantum dots by in situ scanning tunneling probe-assisted nanolithography and molecular beam epitaxy. J. Vac. Sci. Technol., B **2002**, *20* (3), 762–765.
10. Yang, T.; Kohmoto, S.; Nakamura, H.; Asakawa, K. Effects of lateral quantum dot pitch on the formation of vertically aligned InAs site-controlled quantum dots. J. Appl. Phys. **2003**, *93* (2), 1190–1194.
11. Ishikawa, T.; Kohmoto, S.; Asakawa, K. Site control of self-organized InAs dots on GaAs substrates by in situ electron-beam lithography and molecular-beam epitaxy. Appl. Phys. Lett. **1998**, *73* (12), 1712–1714.
12. Kohmoto, S.; Ishikawa, T.; Asakawa, K. InAs-dot/GaAs structures site-controlled by in situ electron-beam lithography and self-organizing molecular beam epitaxy. Jpn. J. Appl. Phys. **1999**, *38*, 1075–1077.
13. Horikoshi, Y.; Yamaguchi, H.; Briones, F.; Kawashima, M. Growth process of III–V compound semiconductors by migration-enhanced epitaxy. J. Cryst. Growth **1990**, *105* (1–4), 326–338.
14. Xie, Q.; Madhukar, A.; Chen, P.; Kobayashi, N. Vertically self-organized InAs quantum box islands on GaAs(100). Phys. Rev. Lett. **1995**, *75* (13), 2542–2545.
15. Solomon, G.S.; Trezza, J.A.; Marshall, A.F.; Harris, J.S., Jr. Vertically aligned and electronically coupled growth induced InAs islands in GaAs. Phys. Rev. Lett. **1996**, *76* (6), 952–955.
16. Springholz, G.; Holy, V.; Pinczolits, M.; Bauer, G. Self-organized growth of three-dimensional quantum-dot crystals with fcc-like stacking and a tunable lattice constant. Science **1998**, *282*, 734–737.
17. Pinczolits, M.; Springholz, G.; Bauer, G. Evolution of hexagonal lateral ordering in strain-symmetrized $PbSe/Pb_{1-x}Eu_xTe$ quantum-dot superlattices. Phys. Rev., B **1999**, *60* (16), 11524–11529.
18. Tersoff, J.; Teichert, C.; Lagally, M.G. Self-organization in growth of quantum dot superlattices. Phys. Rev. Lett. **1996**, *76* (10), 1675–1678.
19. Solomon, G.S.; Komarov, S.; Harris, J.S., Jr.; Yamamoto, Y. Increased size uniformity through vertical quantum dot columns. J. Cryst. Growth **1997**, *175–176*, 707–712.
20. Kohmoto, S.; Nambu, Y.; Asakawa, K.; Ishikawa, T. Reduced nonradiative recombination in etched/regrown AlGaAs/GaAs structures fabricated by in situ processing. J. Vac. Sci. Technol., B **1996**, *14* (6), 3646–3649.
21. Sugimoto, Y.; Ikeda, N.; Carlsson, N.; Asakawa, K.; Kawai, N.; Inoue, K. Fabrication and characterization of different types of two-dimensional AlGaAs photonic crystal slabs. J. Appl. Phys. **2002**, *91* (3), 922–929.
22. Nakamura, H.; Kanamoto, K.; Nakamura, Y.; Ohkouchi, S.; Ikeda, N.; Tanaka, Y.; Sugimoto, Y.; Ishikawa, H.; Asakawa, K. In *Large Enhancement of Optical Nonlinearity Using Quantum Dots Embedded in a Photonic Crystal Structure for All-optical Switch Applications*, Proceedings of LEOS2002, Glasgow, 2002.

Quantum Rods Made of Cadmium Selenide (CdSe): Anistropy

Liang-shi Li
A. Paul Alivisatos
University of California, Berkeley, California, U.S.A.

INTRODUCTION

Semiconductor nanocrystals, also known as "quantum dots," have been intensively studied because of their novel properties caused by quantum confinement and their potential to be used in making electro-optical devices.[1,2] The colloidal nanocrystals synthesized by wet chemistry methods have in particular drawn much attention because of their chemical processibility[3,4] and their capability to be incorporated into various matrices such as conducting polymer[5–8] and biological systems.[9–12] Studies of CdSe nanocrystals with the wurtzite lattice structure have played an important role in our understanding of quantum confinement[13,14] and of the growth mechanism of nanocrystals.[15] The recent advances in synthesizing rodlike,[16–18] disclike,[19,20] and more exotically shaped[21,22] nanocrystals have opened a new field to study the shape dependence of their properties. In this article, we review some of the work on CdSe nanorods (or "quantum rods") that has been done recently in our group. First, we talk about the synthesis of CdSe quantum rods with tightly controlled variable lengths and diameters and second, their optical and dielectric properties. Because the alignment of the quantum rods is desirable for both fundamental studies and applications, we will also cover our work on the formation of the lyotropic nematic phase of these geometrically anisotropic objects in solution, and the large-scale alignment of the quantum rods by taking advantage of the liquid crystalline phase. Then we will finish by briefly discussing some of their potential applications.

GROWTH OF CdSe QUANTUM RODS

Generally, CdSe nanocrystals are made by injecting the mixture of precursors (dimethyl cadmium and Se powder dissolved in tributylphosphine) into a hot surfactant protected in inert gases.[15] The thermal decomposition of the precursors results in the nucleation of the CdSe particles and the subsequent crystal growth. Typically, the precursors are injected at 360°C and then the temperature is lowered to 250°C for crystal growth. When trioctylphosphine oxide (TOPO) is used as the surfactant, the CdSe nanocrystals have nearly spherical shape (aspect ratio <1.2). These nanocrystals are coated by organic molecules so that they do not aggregate easily and can be dispersed in organic solvents. These nanocrystals have wurtzite lattice structure, like CdSe bulk crystals. Recently, Peng et al. found that when a mixture of TOPO and hexylphosphonic acid (HPA) is used as the surfactant, rodlike CdSe nanocrystals can be made.[17] The CdSe nanorods also have a wurtzite crystalline lattice, and are elongated in the *c*-crystallographic axis. However, the surfactant mixture of TOPO and HPA does not give us the capability to control the monodispersity or to tune the length and width of the quantum rods. This is because the growth of the nanorods is so fast that the growth of the nanocrystals is very sensitive to subtle changes in the reaction conditions such as speed of precursor injection. Consequently, nanocrystals with broad size distribution and even with different shapes are made.[21] On the other hand, it was observed that when HPA is replaced with a phosphonic acid with a longer alkyl chain, tetradecylphosphonic acid (TDPA), the speed of the growth of the nanocrystals is dramatically reduced,[23] so much so that only nanorods with very small aspect ratio are made without further addition of the precursors during the growth process. To tune the growth rate of the nanorods, we use a mixture of HPA and TDPA with TOPO, and we found that the controllability of the synthesis is greatly improved.[24,25] By changing the molar ratio of HPA and TDPA, the speed of growth of the nanorods in both transverse and longitudinal directions can be tuned, so that the nanorods with tightly controlled dimensions can be made reproducibly.[24,25] Extraction of part of the reaction solution and subsequent quenching with a large excess of toluene at different time in the early stage of the reaction results in nanorods with different lengths but almost the same diameter. In our study, the HPA/TDPA ratio is systematically changed and the products examined by transmission electron microscopy (TEM). A HPA/TDPA ratio of 1:3 results in nanorods with large diameter (>4 nm) but small aspect ratio (1–10), whereas increasing the amount of HPA in the mixture but with fixed total molar ratio of the phosphonic acids and TOPO produces nanorods with smaller diameter but large aspect ratio. When the HPA/TDPA ratio is higher than 3:1, the reaction loses control and nanoparticles with broad size

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013595

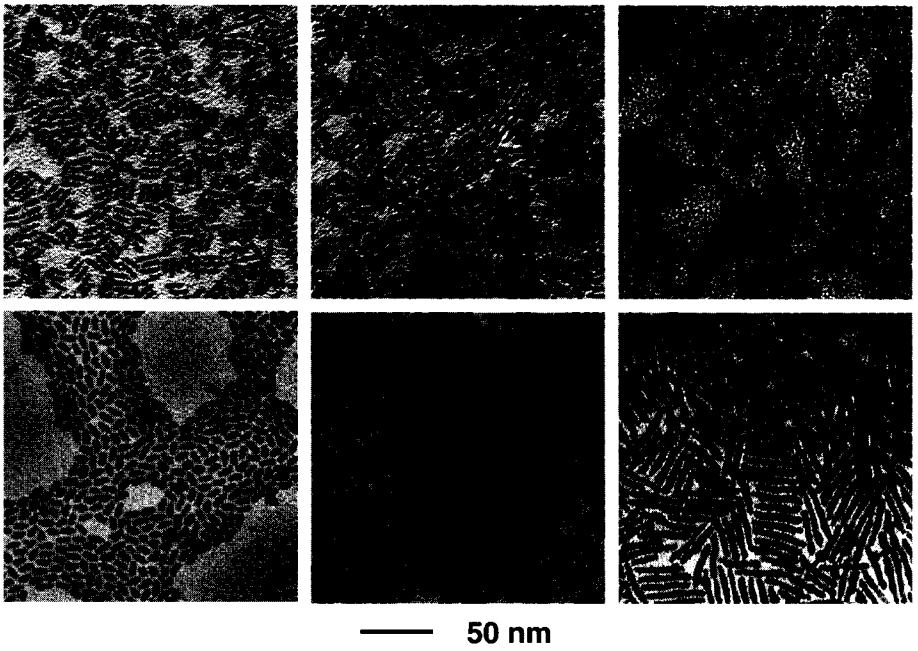


Fig. 1 TEM images of CdSe quantum rods made by using a mixture of HPA, TDPA, and TOPO as the surfactant. By changing the molar ratio of HPA and TDPA, the length and width of the nanocrystals can be systematically controlled. (Reproduced with permission from Ref. [53]. Copyright 2002 American Chemical Society.)

and shape distribution are produced. Fig. 1 show the electron transmission micrographs (TEM) of CdSe quantum rods made with this method. No size selection was performed on the samples after syntheses.

Despite the success in controlling the length and width of CdSe quantum rods, little is known about their microscopic growth mechanism, mainly because of the lack of suitable experimental techniques. However, based on the analysis of the products made by varying the growth conditions, a few observations have been made.[21,23] First, rod growth is possible only when the monomer concentration is high enough. The monomers are believed to be some complexes of cadmium that formed at high temperature and may be very different from the species stable at room temperature, and their nature is not yet known. To maintain the monomer concentration for the quantum rods to steadily grow along the *c* axis, the presence of phosphonic acids is required. Second, when the monomers are depleted, unless more Cd and Se precursors are added the quantum rods will grow fatter and shorter so that eventually nearly spherical nanocrystals are formed. This suggests that the formation of elongated CdSe nanocrystals would be kinetically favored whereas that of nearly spherical ones thermodynamically favored. Third, the growth of the CdSe quantum rods is unidirectional. This is based on the observation that when the amount of HPA in the surfactant mixture is higher than needed for quantum rod growth, nanocrystals with unidirectional shapes such as pencil-, arrow-, or treelike shapes are made.[21]

Based on the observations described above, so far two mechanisms have been proposed: selective adsorption[21] and diffusion-controlled growth mechanisms.[23] The selective adsorption mechanism is based on the different chemical affinity of phosphonic acids to the different faces on the nanocrystals, and thus the different energy of crystal faces in the presence of phosphonic acid molecules. According to this model, all crystal faces except $(00\bar{1})$ is coated by the phosphonic acid molecules, and therefore the nanocrystals grow unidirectionally to form rods. The diffusion-controlled growth mechanism assumes the mass transport of the monomers to be the determining step of the nanocrystal growth, so that the concentration gradient of the monomers in the bulk solution and the stagnant solution around the quantum rods determines the shape evolution of the CdSe nanocrystals.

OPTICAL PROPERTIES

As is well known for (nearly) spherical semiconductor nanocrystals, the absorption and photoluminescence spectra of rodlike nanocrystals are also dependent on their sizes, i.e., both lengths and widths.[25] Fig. 2 shows the length and width dependence of the photoluminescence energy of CdSe quantum rods with various lengths

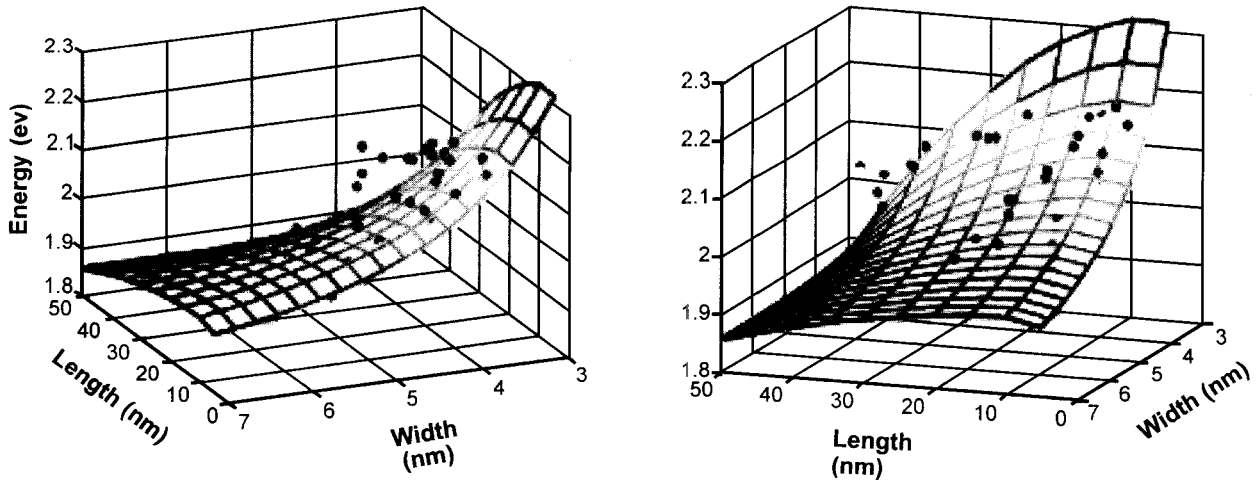


Fig. 2 The photoluminescence energy of CdSe quantum rods as a function of their length and width. It is more sensitive to width than to length. The meshes drawn are the best polynomial fit. (Reproduced with permission from Ref. [25]. Copyright 2001 American Chemical Society.) (*View this art in color at www.dekker.com.*)

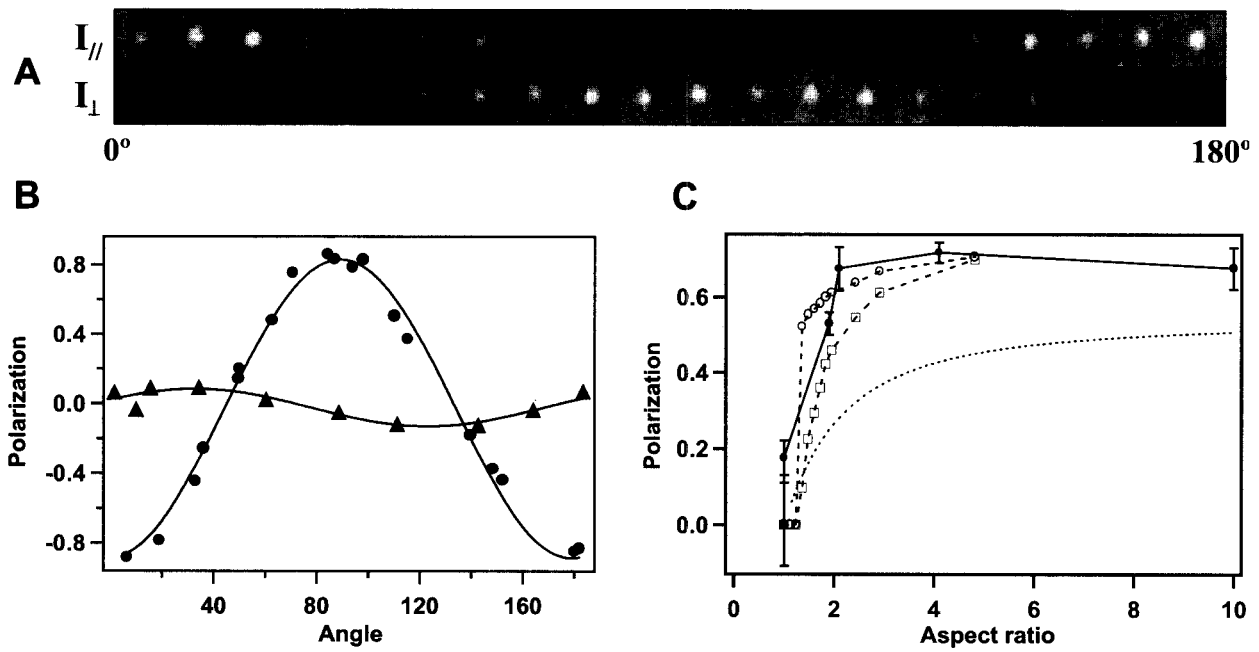


Fig. 3 Luminescence polarization measurements from individual quantum rods at room temperature. (A) Luminescence images of a single rod simultaneously recorded in two perpendicular polarization directions at detection angles changing from 0° to 180°. The rods have an aspect ratio of 10:1. (B) The intensity ratio $r=(I_{\|}-I_{\perp})/(I_{\|}+I_{\perp})$ calculated from the luminescence images in A (filled circles) and fitted with sinusoidal function (solid line) to give a polarization factor of 0.86. The data from spherical dots (filled triangles) show much smaller polarization factors. (C) Polarization factor vs. aspect ratio. The filled circles with error bars are from experiments, unfilled circles from empirical pseudopotential calculations for 0 K, and squares from empirical pseudopotential calculations for room temperature. Each experimental data point was obtained by measuring more than 40 individual rods. The lines simply connect data points, and the dotted line is from the fitting based on the dielectric model proposed in Ref. [27] to the data points. (Reproduced with permission from Ref. [24]. Copyright 2001 AAAS.)

and width. The overall tendency is that with increasing length or width, the luminescence shifts to lower energy, which is the same as that of spherical nanocrystals. These figures also show that the width dependence is stronger than the length dependence in the size range we have studied. This is because the confinement of the photo-excited carriers is mainly imposed in the lateral directions.

It has been well known that elongation in the shape of nanocrystals results in anisotropy in their optical properties.[26,27] Even in the spherical CdSe nanocrystals, polarization of the photoluminescence in the a–b crystal plane has been theoretically predicted[28,29] and experimentally verified,[30] because of crystal field splitting in the hexagonal wurtzite crystal structure. The perturbation treatment of the nearly spherical CdSe nanocrystals[31] suggested that the ellipticity compensated for the crystal field splitting and even changed the order of the lowest electronic states and thus the polarization of the fluorescence. To study the polarization of the photoluminescence emitted from CdSe quantum rods, single-molecule fluorescence spectroscopy was performed on CdSe quantum rods with variable aspect ratio at room temperature.[24] CdSe quantum rods are sparsely deposited on a glass substrate and are excited by a circularly polarized CW Ar^+ laser (480 nm). A beam displacement crystal is used to

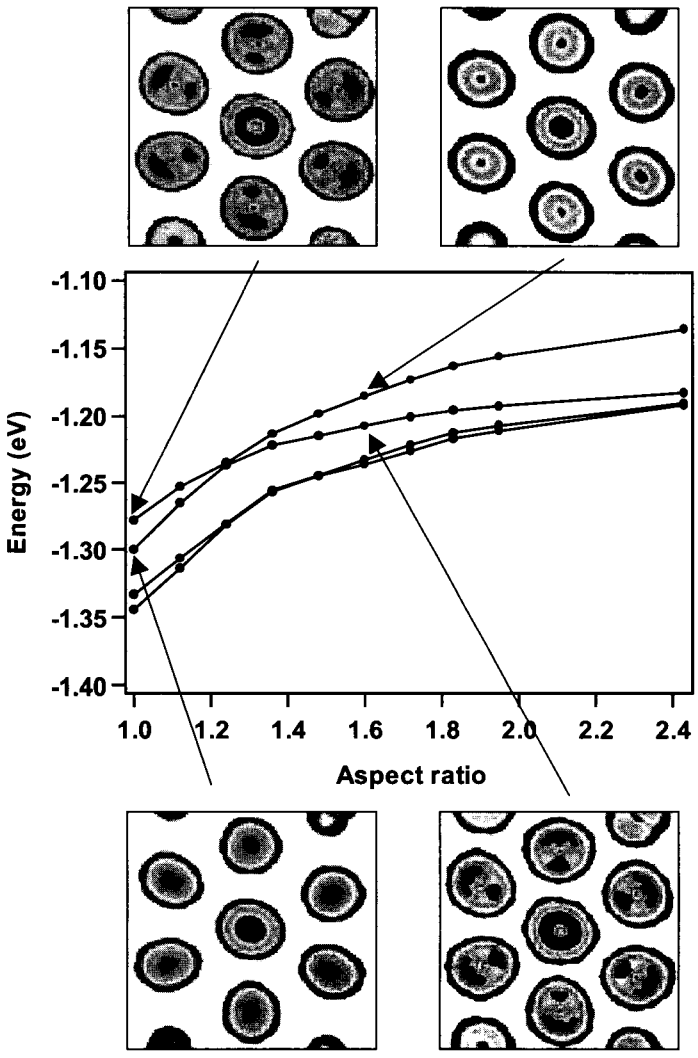


Fig. 4 The four highest occupied electronic states of 3.0-nm-diameter CdSe quantum rods calculated with an empirical pseudopotential method with different aspect ratios. The two highest energy levels have a crossing at an aspect ratio around 1.25. Insets: Contour plots of the two highest occupied states for rods with an aspect ratio of 1.00 and 1.60, respectively. They show the projection of electron density around some Se atoms in a plane, with density increasing from red to blue. This plane intersects the rod in the middle, perpendicular to the long axis (*c* axis). The crossing over of the predominantly $p_{x,y}$ and predominantly p_z levels vs. aspect ratio can be seen. (Reproduced with permission from Ref. [24]. Copyright 2001 AAAS.) (*View this art in color at www.dekker.com.*)

spatially split the photoluminescence from individual quantum rods into two beams with perpendicular polarizations and subsequently imaged on a CCD camera at the same time. The rotation of the crystal and therefore the rotation of the polarization directions of the two split beams show the oscillating intensity. Images of luminescence from a single quantum rod with 10:1 aspect ratio (Fig. 3A) demonstrate the change of the fluorescence intensity with detection angle along the two perpendicular polarization directions. The polarization factor r for this rod sample is determined to be 86% by fitting the intensity ratio $r=(I_{\parallel}-I_{\perp})/(I_{\parallel}+I_{\perp})$ vs. detection angle with a sinusoidal function (filled circle in Fig. 3B). In comparison, the polarization factor of spherical dots is usually less than 10% (filled triangles in Fig. 3B). We have measured samples with aspect ratios from 1:1 to 30:1 at room temperature. The polarization factor changes rapidly from nearly zero to ~70% (e.g., $I_{\parallel}:I_{\perp}\cong 5.5:1$) when aspect ratio increases from 1:1 to 2:1, then remains almost constant afterward (Fig. 3C).

Semiempirical pseudopotential calculation performed on the CdSe quantum rods[24,32] has shown that the linearly polarized photoluminescence from the CdSe quantum rods is indeed a result of the energy level crossing in the highest occupied electronic states. Qualitatively, the lowest excited electronic state of the bulk CdSe semiconductor arises when an electron is excited from an occupied Se 4p atomic orbital to an empty Cd 5s orbital. This excitation is distributed throughout many unit cells of the solid, and there is a pronounced dependence of the bandgap on the size of the crystals. Because the p atomic orbitals are degenerate, the interaction between Se 4p orbitals in the crystal field and the effect of spin–orbit coupling play important roles in the electronic structure of CdSe nanocrystals. The empirical pseudopotential calculation was performed to study the evolution of the electronic structure when CdSe nanocrystals evolve from a spherical to a rodlike shape. We started from a spherical dot with a diameter of 3.0 nm and increased the aspect ratio by inserting a cylindrical segment along the c axis. The four lowest unoccupied and four highest occupied electronic states were calculated for quantum rods with aspect ratios ranging from 1:1 to 5:1. As the aspect ratio increases from unity, a crossover of the two highest occupied electronic states occurs at an aspect ratio of 1.25 (Fig. 4). Before this point, the electronic states with predominantly Se $4p_x$, $4p_y$ components have higher energy than that of states that are mainly $4p_z$. The electronic energy levels of these states all increase with increasing size. The $4p_z$ orbital has greater momentum projected onto the c axis of the crystal compared to $4p_x$ and $4p_y$, so the energy levels with greater $4p_z$ component are more sensitive to the rod length. These states exchange position relative to each other at an aspect ratio greater than 1.25. Calculations for the cases of 2.1- and 3.8-nm-wide rods show similar results and the crossing point is 1.25 and 1.36, respectively.

DIELECTRIC PROPERTIES

In contrast to their well-studied optical and electronic properties, the dielectric properties of CdSe nanocrystals have been seldom investigated, especially their permanent dipole moment. On one hand, theoretical analyses have suggested the presence of a permanent dipole moment along the c crystallographic axis because of the lack of inversion symmetry of the wurtzite structure;[33–35] on the other hand, dielectric dispersion measurements performed on concentrated solutions of the spherical CdSe[36,37] nanocrystals have revealed a permanent dipole moment as a result of surface trapped charges, while the contribution of the noncentrosymmetric lattice is negligible in the size range studied.

However, the surface charge model does not address the direction of the permanent dipole moment, nor does the dielectric dispersion measurement itself because it does not distinguish the uniaxial nature of the spherical nanocrystals. The elongated shape of the CdSe nanorods gives us an opportunity to answer this question.

Transient electric birefringence (TEB) is a method that has been widely used to study the rotational diffusion, size, shape, and polarization properties of objects with anisotropic geometry, especially macromolecules and biological systems such as DNA, viruses, and proteins.[38,39] The transient behavior of the birefringence reflects the alignment mechanism of these molecules in response to a pulsed electric field. In particular, for a suspension of elongated objects with axial symmetry and a large aspect ratio (≥ 5), the rising (Δn_r) and falling (Δn_f) edges of the birefringence (difference between the refractive indices of the sample along and perpendicular to the applied electric field) on application and removal of an external electric field are given, respectively, by[40]

$$\Delta n_r(t) = \Delta n_s\left(1+\frac{\gamma-2}{2(\gamma+1)}\exp(-6D_R t) - \frac{3\gamma}{2(\gamma+1)}\exp(-2D_R t)\right)$$

$$\Delta n_f(t) = \Delta n_s \exp(-6D_R t) \quad (1)$$

where $\Delta n_s = \Delta n_r(t\rightarrow\infty)$, $\gamma = \mu_z'^2/[k_b T(\alpha_{\parallel}-\alpha_{\perp})]$, μ'_z is the screened value of the electric dipole moment along the long axis of the rods, $\alpha_{\parallel}$ and $\alpha_{\perp}$ the static electric polarizability along and perpendicular to the long axis, respectively, and D_R the rotational diffusion constant

around an axis normal to the long axis of the rods. According to Eq. 1, the rise time of the birefringence will be slower than the fall time only when there is a permanent dipole moment along the long axes of the nanorods. Benoit[40] used this method to study tobacco mosaic virus (TMV), and from the symmetric falling and rising edges he concluded that TMVs do not have permanent dipole moment in aqueous solution.

To measure the electric birefringence we use the procedure of O'Konski and Zimm,[41,42] as described in detail in Ref. [43]. The quantum rods under electric field have birefringence of the same sign as nitrobenzene, indicating that the nanorods align along (rather than perpendicular to) the electric field. Fig. 5A shows a typical transient electric birefringence curve measured for CdSe nanorods.[43] As suggested by Eq. 1, the asymmetric

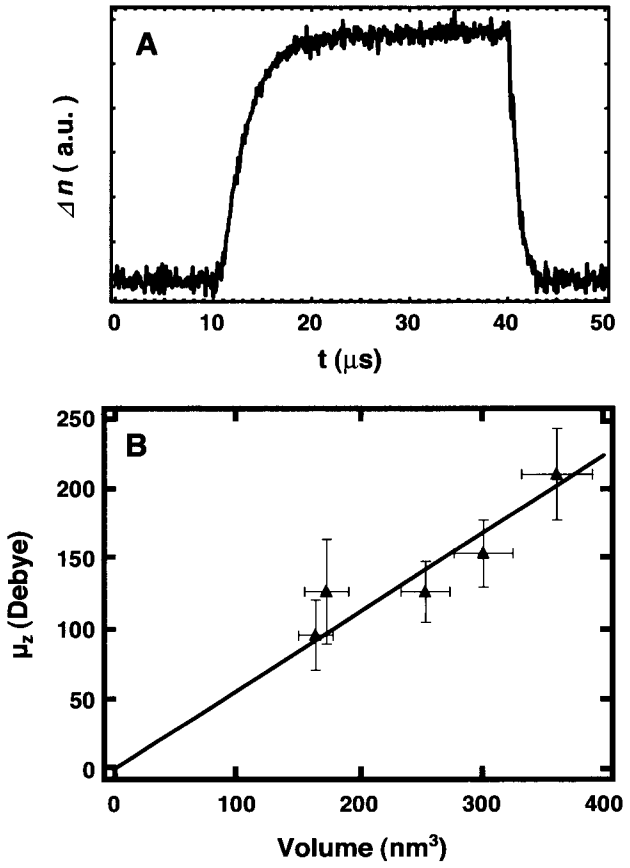


Fig. 5 (A) A typical TEB curve (in linear scale) on the application and subsequent removal of the electric field. The asymmetric rising and falling edges indicate the existence of a permanent dipole moment along the long axis of the quantum rods. In this particular case, the CdSe quantum rods are 4.8 nm wide and 30 nm long. (B) The unscreened dipole moment of CdSe quantum rods with different sizes. The solid triangles are the values measured with the TEB method, and the straight line is the best fit, with a slope of 0.19 μC/cm^2. (Modified with permission from Ref. [43]. Copyright 2003 American Physical Society.)

falling and rising edges indicate the existence of a permanent electric dipole moment along the long axis. The hexagonal symmetry of CdSe nanorods allows us to approximately treat them as being axially symmetric, thus we can fit the TEB curves with Eq. 1 to obtain the rotational diffusion constants and the ratio of permanent dipole moment to the polarizability anisotropy (γ).

To get the permanent dipole moment from γ, we calculate the electric polarizability of CdSe nanorods by assuming a revolute prolate shape for the nanorods, so that the principle axes of polarizability coincide with the geometrical axes. The polarizabilities are calculated by[44]

$$\alpha_{\|(\perp)} = \varepsilon_0 v(\varepsilon_{\|(\perp)} - 1)/[1 + A_{\|(\perp)}(\varepsilon_{\|(\perp)} - 1)]$$

where v is the volume of individual nanorod, ε's the relative dielectric constants between nanorods and the solvent along ($\|$) or perpendicular ($\perp$) to the long axis of the nanorods, and A's the geometrical factors that can be calculated from the dimension of the nanorods.[45] The dielectric constants of CdSe nanorods are taken as bulk material values ($\varepsilon_{\|}$=10.2, $\varepsilon_{\perp}$=9.33), and that of cyclohexane is taken as 2.02.

The unscreened permanent dipole moments μ_z of five nanorod samples are plotted in Fig. 5B vs. their volume. Considering the ensemble nature of the measurement, these are the root mean square dipole moment of the samples. It is an average not only over the finite size distribution, but also over the possible structural distribution. An empirical pseudopotential calculation showed that the dipole moment strongly depends on the detailed structure of the nanocrystals and can vary significantly on small structural changes.[35] Within the experimental error, however, a linear dependence of μ_z vs. volume is obtained, and no correlation between rod length or width and μ_z can be realized. This is consistent with theoretical analyses that the polarity is intrinsic to the crystallographic lattice of CdSe because of the lack of inversion symmetry. Our results are not consistent with a random dipole moment resulting from trapped surface charges, as proposed for spherical CdSe nanocrystals. From the slope, we get the polarization of CdSe nanorods to be 0.19 μC/cm^2 along the c crystallographic axis, which is in good agreement with the value of 0.6 μC/cm^2 as estimated[34] from a phenomenological rule that was proved experimentally only for ferroelectric materials.

LYOTROPIC LIQUID CRYSTALLINE PHASE OF CdSe QUANTUM ROD DISPERSION

The anisotropic properties of the quantum rods may find applications in devices such as solar cells, polarizing light-emitting diodes, and electro-optical switches. However, to

achieve this, it is important to align the quantum rods on a large scale. Recently, enormous effort has been devoted to aligning the geometrically anisotropic nanoscale objects such as nanowires, nanotubes, and nanorods.[16,46] On the other hand, it has been well known that rodlike or disclike objects, when dispersed in a solvent to a high enough concentration, will spontaneously form lyotropic liquid crystalline phases in which the geometrically anisotropic objects are orientationally ordered. This is clearly illustrated by the rigid rod model proposed by Onsager in 1949.[47] In this model, the free energy of a dispersion of long, thin spherocylinders (aspect ratio > 100) in a solvent can be obtained by a virial expansion,[48]

$$\frac{\Delta F}{NkT} = \frac{\mu^0}{kT} + \ln(\Lambda^3\rho) - 1 + \left\{ \int f(\Omega)\ \ln[4\pi f(\Omega)]\ \mathrm{d}\Omega + \frac{1}{2}\rho \iint \upsilon_{\text{exclusive}}(\Omega, \Omega') f(\Omega) f(\Omega')\mathrm{d}\Omega\ \mathrm{d}\Omega' + \cdots \right\}$$

where $f(\Omega)$ is the orientational distribution function of the rodlike objects, and ρ the number density of the spherocylinders. The first term in the curved bracket is called orientational entropy, which favors a random orientational distribution; while the second term is called packing entropy, which favors an orientationally ordered phase. The packing entropy term dominates only when the density of the rods is high enough so that a nematic lyotropic phase can spontaneously form. Famous examples include rodlike organic macromolecules such as poly(γ-benzyl α,L-glutamate) (PBLG)[49] rodlike tobacco mosaic virus,[50] etc. Lyotropic liquid crystalline mesophases based on all-inorganic minerals have also been known for quite a long time, and have been reviewed by several authors;[48,51,52] however, this area has rarely been explored because of the lack of the interest other than purely academic purposes, in contrast with their thermotropic counterpart. The recent success in the synthesis of the highly soluble and processible metallic semiconductors and magnetic nanocrystals with anisotropic geometries[16–20] and the need for aligning them in order to use their anisotropic properties have provided us a new opportunity to study these liquid crystalline phases.

The most important requirement for the formation of lyotropic liquid crystalline phases is the high dispersion concentration, which is not trivial for most of the inorganic nanocrystals. Recently, we have succeeded in making the nematic phase of solution of CdSe quantum rods in organic solvent.[53] The solvent we use is anhydrous cyclohexane or anhydrous hexanes, which can dissolve CdSe quantum rods to a volume percentage up to 60%. Dilute solutions of quantum rods in a vial are concentrated by blowing with dry nitrogen, and a birefringent phase starts to appear at a certain concentration as observed under an optical polarizing microscope. The solutions are then transferred to NMR tubes or capillary tubes for further study. Special care has to be taken during the process to avoid the exposure of the solution to the air, because the inclusion of even trace amount of water would result in thermoreversible gelation,[54–56] which is also often seen in suspensions of organic rodlike polymers and has been attributed to the degradation of the solvent. Fig. 6A shows isotropic-liquid crystalline phase separation in a 300-μm-diameter capillary tube, and the characteristic disclinations for a nematic phase observed

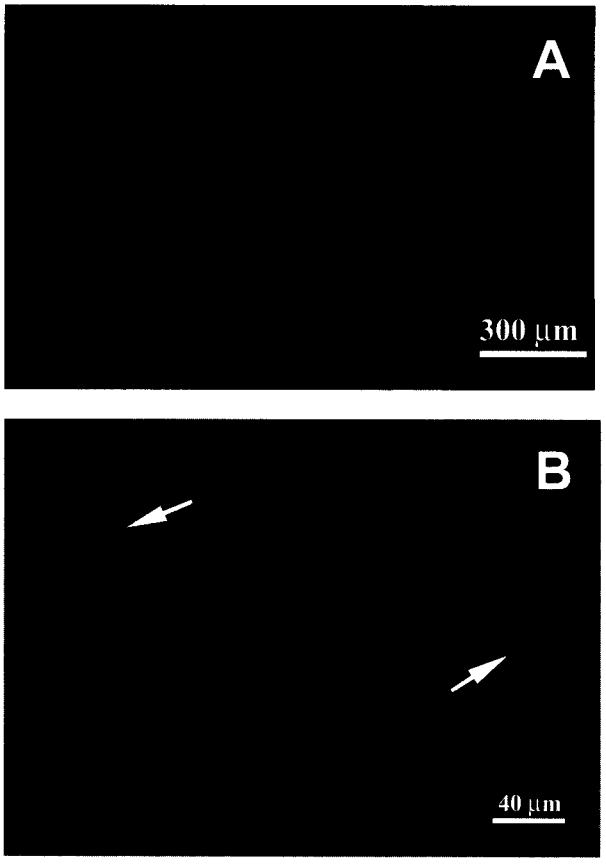


Fig. 6 (A) Isotropic-liquid crystalline phase separation observed between crossed polarizers. The birefringent liquid crystalline phase is on the bottom due to the high density, while the isotropic phase on the top is not visible because of the complete extinction of the transmitted light. (B) The typical defect structures observed in a thin film of liquid crystalline solution of CdSe nanorod solution. The arrows point to the disclinations. The red color is due to the absorption of the CdSe nanocrystals. (Modified with permission from Ref. [53]. Copyright 2002 American Chemical Society.) (*View this art in color at www.dekker.com.*)

in a thin layer of the solution of CdSe 3.8×40-nm quantum rods. The red color is due to the absorption of this quantum rod sample.

The formation of the liquid crystalline phase allows us to align the quantum rods by taking advantage of the collective behavior in a liquid crystalline solution.[57] Various methods have been applied to align the liquid crystalline materials such as electric and magnetic fields as well as curved surfaces. Fig. 7 shows the X-ray diffraction patterns of a liquid crystalline solution of CdSe quantum rods in a 300 μm capillary tube.[58] These rods have a length of ~60 nm and width of 3.0 nm. In both figures, the capillary tube remains vertical. In Fig. 7A, the sharp arcs in the vertical position are due to (002) planes of the CdSe crystalline lattice within individual quantum rods. The high intensity and sharpness of the arcs are determined by the elongated shape of the CdSe quantum rods. The anisotropy of the pattern demonstrates macroscopic alignment of the long axis of the rods. Consistently, in Fig. 7B the diffuse arcs in the equator direction correspond to the lateral distance of the quantum rods, suggesting the short axis is aligned as well. Together, the patterns demonstrate the preferential alignment of the CdSe quantum rods along the capillary tube axis direction, which is attributed to the curved surface of the capillary tube. It can be estimated that about 70% of the quantum rods are oriented within 20° from the axis of the capillary tube.

Alignment of lyotropic liquid crystalline solution with magnetic fields has been reported.[59] Even when none of the components is ferromagnetic, the anisotropy of the magnetic susceptibility of individual constituent molecules adds up because the long-range orientational correlation in the liquid crystalline phases, and therefore the energy gain because of the alignment can easily exceed the thermal fluctuation kT. The alignment of CdSe quantum rods in a liquid crystalline solution under a magnetic field is observed with NMR. A few drops of deuterated chloroform are added into a solution of CdSe quantum rods in cyclohexane, which has a high enough concentration to have both isotropic and liquid crystalline phases present. Fig. 8 shows our preliminary result of DNMR spectrum observed in a 500-MHz spectrometer at 30°C. The central peak is confirmed to be due to the isotropic part of the solution, while the other two peaks are due to the quadruple splitting of the deuterium nuclei in the anisotropic liquid crystalline medium. The comparison of the spectrum with a typical Pake powder pattern[60,61] indicates the alignment of the C–D bond in $CDCl_3$, presumably as a result of both the alignment of the quantum rods by the magnetic field and the interaction between the nanorods and $CDCl_3$ molecules. The exact mechanism is still under investigation.

In spite of the formation of the liquid crystalline phase of the CdSe quantum rod solution, tremendous challenges exist in the alignment of the CdSe quantum rods, especially on a substrate. It is not straightforward to apply the techniques used in liquid crystal display industry such as rubbed surfaces.[62] Furthermore, the volatile nature of the solvent in the CdSe liquid crystalline solution and the propensity of the solution to gel on the exposure to the air or water make the solution difficult to handle.

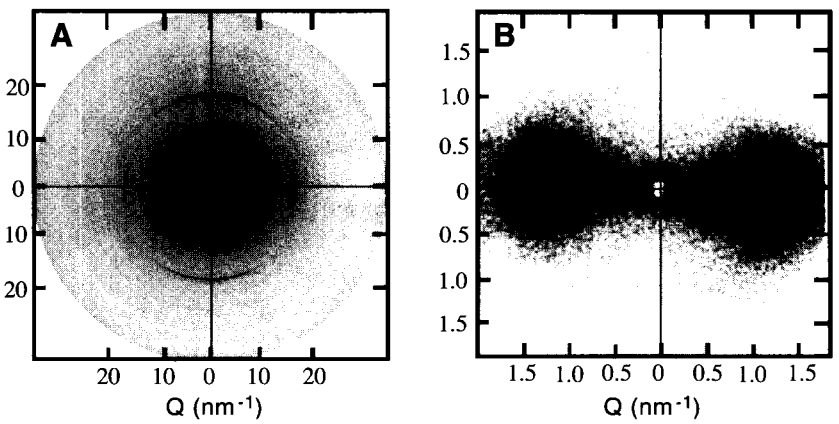


Fig. 7 Wide-angle X-ray diffraction and small-angle X-ray scattering of the same nematic solution of CdSe quantum rods in a 300-μm-thick capillary tube show the preferential alignment of these quantum rods along the axis of the capillary tube. In both measurements, the capillary tube is held vertically. (A) Wide-angle X-ray diffraction pattern. The vertical sharp, intense arcs correspond to the (002) planes of wurtzite CdSe lattice. At about the same radius, weak arcs corresponding to the (100) and (101) planes can also be seen, but they are much more diffuse. (B) Small-angle X-ray scattering pattern. The two diffuse arcs at $Q \sim 1.5$ nm^{-1} on the equator corresponds to the lateral spacing between the rods in the nematic phase. The arcs corresponding to the longitudinal spacing are not resolved. (Reproduced with permission from Ref. [58]. Copyright 2003 Wiley-VCH.)

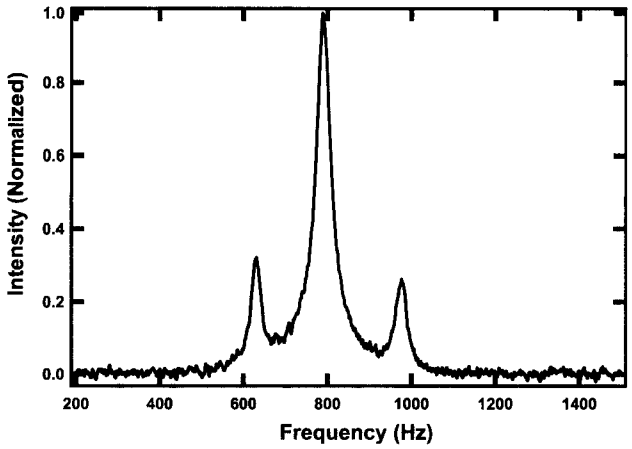


Fig. 8 The deuterium NMR spectrum of a solution of CdSe quantum rods at 30°C in which isotropic phase and liquid crystalline phase coexist. The solution is doped with $CDCl_3$. The central peak is caused by the isotropic part of the solution, which significantly broadens because of the high viscosity of the solution, while the two side peaks are caused by quadruple splitting of the deuterium nuclei in the liquid crystalline phase. The comparison with a powder pattern indicates the alignment of the quantum rods in the liquid crystalline solution.

POTENTIAL APPLICATIONS

Even though the development of the devices with semiconductor quantum rods is still in the early stage, the recent success in using CdSe quantum rods–conjugated polymer composite to make solar cell with high power conversion efficiency[63] has demonstrated the huge potential of the nanocrystals as an active component in electro-optical devices. The solar cells made with inorganic–organic composite take advantage of both the high carrier mobility in inorganic semiconductors and the easy processibility of organic polymers. Furthermore, the presence of the interface between two materials with very different electron affinities enhances the charge separation of the photo-excited carriers. The high extinction coefficient of CdSe nanocrystals makes a very thin film ($\sim$100 nm) of the composite adequate to absorb all of the solar radiation. The application of CdSe quantum rods in the composite, as compared to the nearly spherical nanocrystals, provides a natural path for electrical transport after the charge separation, so that the probability of recombination of photocarriers is reduced. Recently, a power conversion efficiency of 1.7% has been achieved by mixing 7×60-nm CdSe quantum rods with poly(3-hexylthiophene) (P3HT). The rod length dependence study shows that the charge transport improves substantially when the aspect ratio of the quantum rods increases from 1 to 10. Further improvement to the power conversion efficiency is still on the way, which includes the alignment of the quantum rods perpendicular to the film, the improvement to the interface between the nanocrystals and the polymer, and using nanocrystals of other materials such as CdTe. Other possible applications include polarizing light-emitting diodes and electro-optical switches by taking advantage of the anisotropic linear and nonlinear optical properties of these quantum rods.

CONCLUSION

In summary, we have discussed the recent development in the study of CdSe quantum rods. It is just the beginning of the investigation of the shape-dependent properties of semiconductor nanocrystals, and we believe the advent of the new methods to make nanocrystals with more complicated shapes, such as tetrapods and even dendritic inorganic nanocrystals,[21,22] will further advance this kind of study. In addition, the success in the macroscopic alignment of the nanorods provides us unique opportunities to investigate some anisotropic ensemble properties such as optical or X-ray absorption, surface selectivity of ligands, interaction between nanorods and solvent, etc. However, much work still needs to be done to make the nanorods applicable for devices. Currently work is under way in our laboratory to make the liquid crystalline solutions more robust so that they can survive various processing procedures. Furthermore, methods to align the nanorods on substrates have to be developed.

REFERENCES

1. Alivisatos, A.P. Semiconductor clusters, nanocrystals, and quantum dots. Science **1996**, *271* (5251), 933–937.
2. Brus, L. Quantum crystallites and nonlinear optics. Appl. Phys., A **1991**, *53* (6), 465–474.
3. Steigerwald, M.L.; Alivisatos, A.P.; Gibson, J.M.; Harris, T.D.; Kortan, R.; Muller, A.J.; Thayer, A.M.; Duncan, T.M.; Douglass, D.C.; Brus, L.E. Surface derivatization and isolation of semiconductor cluster molecules. J. Am. Chem. Soc. **1988**, *110* (10), 3046–3050.
4. Schmid, G. Large clusters and colloids—Metals in the embryonic state. Chem. Rev. **1992**, *92* (8), 1709–1727.
5. Colvin, V.L.; Schlamp, M.C.; Alivisatos, A.P. Light-emitting-diodes made from cadmium selenide nanocrystals and a semiconducting polymer. Nature **1994**, *370* (6488), 354–357.
6. Dabbousi, B.O.; Bawendi, M.G.; Onotsuka, O.;

Rubner, M.F. Electroluminescence from CdSe quantum-dot polymer composites. Appl. Phys. Lett. **1995**, *66* (11), 1316–1318.
7. Huynh, W.U.; Peng, X.; Alivisatos, A.P. CdSe nanocrystal rods/poly(3-hexylthiophene) composite photovoltaic devices. Adv. Mater. **1999**, *11* (11), 923–927.
8. Tessler, N.; Medvedev, V.; Kazes, M.; Kan, S.H.; Banin, U. Efficient near-infrared polymer nanocrystat light-emitting diodes. Science **2002**, *295* (5559), 1506–1509.
9. Bruchez, M.; Moronne, M.; Gin, P.; Weiss, S.; Alivisatos, A.P. Semiconductor nanocrystals as fluorescent biological labels. Science **1998**, *281* (5385), 2013–2016.
10. Chan, W.C.W.; Nie, S.M. Quantum dot bioconjugates for ultrasensitive nonisotopic detection. Science **1998**, *281* (5385), 2016–2019.
11. Parak, W.J.; Boudreau, R.; Le Gros, M.; Gerion, D.; Zanchet, D.; Micheel, C.M.; Williams, S.C.; Alivisatos, A.P. Cell motility and metastatic potential studies based on quantum dot imaging of phagokinetic tracks. Adv. Mater. **2002**, *14* (12), 882–885.
12. Loweth, C.J.; Caldwell, W.B.; Peng, X.; Alivisatos, A.P.; Schultz, P.G. DNA-based assembly of gold nanocrystals. Angew. Chem., Int. Ed. Engl. **1999**, *38* (12), 1808–1812.
13. Efros, A.L.; Rosen, M. The electronic structure of semiconductor nanocrystals. Annu. Rev. Mater. Sci. **2000**, *30*, 475–521.
14. Nirmal, M.; Norris, D.J.; Bawndi, M.G.; Efros, A.L.; Rosen, M. Observation of the dark exciton in CdSe quantum dots. Phys. Rev. Lett. **1995**, *75* (20), 3728–3781.
15. Peng, X.; Wickham, J.; Alivisatos, A.P. Kinetics of II–VI and III–V colloidal semiconductor nanocrystal growth: "Focusing" of size distributions. J. Am. Chem. Soc. **1998**, *120* (21), 5343–5344.
16. Li, M.; Schnablegger, H.; Mann, S. Coupled synthesis and self-assembly of nanoparticles to give structures with controlled organization. Nature **1999**, *402* (6760), 393–395.
17. Peng, X.; Manna, L.; Yang, W.; Wickham, J.; Scher, E.; Kadavanich, A.; Alivisatos, A.P. Shape control of CdSe nanocrystals. Nature **2000**, *404* (6773), 59–61.
18. Chang, S.; Shih, C.; Chen, C.; Lai, W.; Wang, C.R.C. The shape transition of gold nanorods. Langmuir **1999**, *15* (3), 701–709.
19. Puntes, V.F.; Zanchet, D.; Erdonmez, C.K.; Alivisatos, A.P. Synthesis of hcp-Co nanodisks. J. Am. Chem. Soc. **2002**, *124* (43), 12874–12880.
20. Jin, R.C.; Cao, Y.W.; Mirkin, C.A.; Kelly, K.L.; Schatz, G.C.; Zheng, J.G. Photoinduced conversion of silver nanospheres to nanoprisms. Science **2001**, *294* (5548), 1901–1904.
21. Manna, L.; Scher, E.C.; Alivisatos, A.P. Synthesis of soluble and processable rod-, arrow-, teardrop-, and tetrapod-shaped CdSe nanocrystals. J. Am. Chem. Soc. **2000**, *122* (51), 12700–12706.
22. Jun, Y.; Lee, S.; Kang, N.; Cheon, J. Controlled synthesis of multi-armed CdS nanorod architectures using monosurfactant system. J. Am. Chem. Soc. **2001**, *123* (21), 5150–5151.
23. Peng, Z.A.; Peng, X. Mechanisms of the shape evolution of CdSe nanocrystals. J. Am. Chem. Soc. **2001**, *123* (7), 1389–1395.
24. Hu, J.; Li, L.-S.; Yang, W.; Manna, L.; Wang, L.-W.; Alivisatos, A.P. Linearly polarized emission from colloidal semiconductor quantum rods. Science **2001**, *292* (5524), 2060–2063.
25. Li, L.-S.; Hu, J.; Yang, W.; Alivisatos, A.P. Band gap variation of size- and shape-controlled colloidal CdSe quantum rods. Nano Lett. **2001**, *1* (7), 349–351.
26. El-Sayed, M.A. Some interesting properties of metals confined in time and nanometer space of different shapes. Acc. Chem. Res. **2001**, *34* (4), 257–264.
27. Kovalev, D.; Benchorin, M.; Diener, J.; Koch, F.; Efros, A.L.; Rosen, M.; Gippius, N.A.; Tikhodeev, S.G. Porous Si anisotropy from photoluminescence polarization. Appl. Phys. Lett. **1995**, *67* (11), 1585–1587.
28. Efros, A.L. Luminescence polarization of CdSe microcrystals. Phys. Rev., B **1992**, *46* (12), 7448–7458.
29. Efros, A.L.; Rosen, M.; Kuno, M.; Nirmal, M.; Norris, D.J.; Bawendi, M. Band-edge exciton in quantum dots of semiconductors with a degenerate valence band: Dark and bright exciton states. Phys. Rev., B **1996**, *54* (7), 4843–4856.
30. Empedocles, S.A.; Neuhauser, R.; Bawendi, M.G. Three-dimensional orientation measurements of symmetric single chromophores using polarization microscopy. Nature **1999**, *399* (6732), 126–130.
31. Efros, A.L.; Rodina, A.V. Band-edge absorption and luminescence of nonspherical nanometer-size crystals. Phys. Rev., B **1993**, *47* (15), 10005–10007.
32. Hu, J.; Wang, L.-W.; Li, L.-S.; Yang, W.; Alvisatos, A.P. Semiempirical pseudopotential calculation of electronic states of CdSe quantum rods. J. Phys. Chem. **2002**, *106* (10), 2447–2452.
33. Huong, N.Q.; Birman, J.L. Origin of polarization in polar nanocrystals. J. Chem. Phys. **1998**, *108* (5), 1769–1772.
34. Schmidt, M.E.; Blanton, S.A.; Hines, M.A.; Guyot-Sionnest, P. Polar CdSe nanocrystals: Implications

for electronic structure. J. Chem. Phys. **1997**, *106* (12), 5254–5259.
35. Rabani, E.; Hetényi, B.; Berne, B.J.; Brus, L.E. Electronic properties of CdSe nanocrystals in the absence and presence of a dielectric medium. J. Chem. Phys. **1999**, *110* (11), 5355–5369.
36. Blanton, S.A.; Leheny, R.L.; Hines, M.A.; Guyot-Sionnest, P. Dielectric dispersion measurements of CdSe nanocrystal colloids: Observation of a permanent dipole moment. Phys. Rev. Lett. **1997**, *79* (5), 865–868.
37. Shim, M.; Guyot-Sionnest, P. Permanent dipole moment and charges in colloidal semiconductor quantum dots. J. Chem. Phys. **1999**, *111* (15), 6955–6964.
38. O'Konski, C.T. Kerr Effect. In *Encyclopedia of Polymer Science and Technology*; Interscience: New York, 1968; Vol. 9, 551–590.
39. Yoshioka, K.; Watanabe, H. Dielectric Properties of Proteins II. Electric Birefringence and Dichroism. In *Physical Principles and Techniques of Protein Chemistry*; Leach, S.J., Ed.; Academic Press: New York, 1969; 335–367. Pt. A.
40. Benoit, H. Contribution a L'Étude de L'Effet Kerr Présenté par Les Solutions Dilutées de Macromolecules Rigides. Ann. Phys. (Paris) **1951**, *6*, 561–608.
41. O'Konski, C.T.; Zimm, B.H. New method for studying electrical orientation and relaxation effects in aqueous colloids: Preliminary results with tobacco mosaic virus. Science **1950**, *111* (2875), 113–116.
42. Newman, J.; Swinney, H.L. Length and dipole moment of TMV by laser signal-averaging transient electric birefringence. Biopolymers **1976**, *15*, 301–315.
43. Li, L.-S.; Alivisatos, A.P. Origin and scaling of the permanent dipole moment in CdSe nanorods. Phys. Rev. Lett. **2003**, *90* (9), 097402.
44. Böttcher, C.J.F. *Theory of Electric Polarization*; Elsevier Scientific: New York, 1973.
45. Osborn, J.A. Demagnetizing factors of the general ellipsoid. Phys. Rev. **1945**, *67* (11), 351–357.
46. Kim, F.; Kwan, S.; Akana, J.; Yang, P.D. Langmuir–Blodgett nanorod assembly. J. Am. Chem. Soc. **2001**, *123* (18), 4360–4361.
47. Onsager, I. The effect of shape on the interaction of colloidal particles. Ann. N. Y. Acad. Sci. **1949**, *51* (4), 627–659.
48. Vroege, G.J.; Lekkerkerker, H.N.W. Phase transitions in lyotropic colloidal and polymer liquid crystals. Rep. Prog. Phys. **1992**, *55* (8), 1241–1309.
49. Robinson, C.; Ward, J.C. Liquid-crystalline structures in polypeptides. Nature **1957**, *180* (4596), 1183–1184.
50. Bawden, F.C.; Pirie, N.W.; Bernal, J.D.; Fankuchen, I. Liquid crystalline substances from virus-infected plants. Nature **1936**, *138*, 1051–1052.
51. Gabriel, J.-C.P.; Davidson, P. New trends in colloidal liquid crystals based on mineral moieties. Adv. Mater. **2000**, *12* (1), 9–20.
52. Sonin, A.S. Inorganic lyotropic liquid crystals. J. Mater. Chem. **1998**, *8* (12), 2557–2574.
53. Li, L.-S.; Walda, J.; Manna, L.; Alivisatos, A.P. Semiconductor nanorod liquid crystals. Nano Lett. **2002**, *2* (6), 557–560.
54. Gunton, J.D.; San Miguel, M.; Sahni, P.S. The Dynamics of First Order Phase Transition. In *Phase Transitions and Critical Phenomena*; Domb, C., Lebowitz, J.L., Eds.; Academic Press: New York, 1983; Vol. 8, 269–482.
55. Tohyama, K.; Miller, W.G. Network structure in gels of rod-like polypeptides. Nature **1981**, *289* (5800), 813–816.
56. Frey, M.W.; Cuculo, J.A.; Ciferri, A.; Theil, M.H. A review of lattice theory for lyotropic liquid crystalline polymers, spinodal decomposition, and gel formation. J. M. S. Rev. Macromol. Chem. Phys., C **1995**, *35*, 287–325.
57. Chandrasekhar, S. *Liquid Crystals*; University Press: Cambridge, 1992.
58. Li, L.-S.; Alivisatos, A.P. Semiconductor nanorod liquid crystals and their assembly on a substrate. Adv. Mater. **2003**, *15* (5), 408–411.
59. Firouzi, A.; Schaefer, D.J.; Tolbert, S.H.; Stucky, G.D.; Chmelka, B.F. Magnetic-field-induced orientational ordering of alkaline lyotropic silicate-surfactant liquid crystals. J. Am. Chem. Soc. **1997**, *119* (40), 9466–9477.
60. Forrest, B.J.; Reeves, L.W. New lyotropic liquid crystals composed of finite nonspherical micelles. Chem. Rev. **1981**, *81* (1), 1–14.
61. Seelig, J. Deuterium magnetic resonance: Theory and application to lipid membranes. Q. Rev. Biophys. **1977**, *10* (3), 353–418.
62. Jerome, B. Surface effects and anchoring in liquid crystals. Rep. Prog. Phys. **1991**, *54* (3), 391–451.
63. Huynh, W.U.; Dittmer, J.J.; Alivisatos, A.P. Hybrid nanorod–polymer solar cells. Science **2002**, *295* (5564), 2425–2427.

Raman Spectroscopy Studies of Carbon Nanotube-Polymer Composites

Bin Chen
National Aeronautics and Space Administration (NASA), Moffett Field, California, U.S.A.

INTRODUCTION

Natural carbon exhibits its distinct properties in two forms: diamond, an insulator, gives the strongest mechanical strength; graphite, a brittle material, conducts electricity at room temperature. A recently synthesized carbon form, the carbon nanotube (CNT), combines the mechanical strength of diamonds with the electric conductivity of graphite. However, the as-grown form of this almost perfect new material is actually a mixture of semiconducting and metallic tubes, in the approximate 2:1 ratio estimated by the tight binding model.[1] Furthermore, as-grown single-wall nanotubes (SWNTs) generally possess a very high aspect ratio of length to diameter (on the order of 1000:1). They tend to bundle together as ropelike hydrophobic aggregates—not an individual tube of either semiconductor or metallic tube. However, controlling growth and follow-up processing and treatment conditions can change the metallic and semiconducting ratio in the mixture (see "Carbon Nanotubes").

To take advantage of their unique electronic properties, mixed SWNT bundles can be dispersed and separated to tailor to various device applications. One ideal matrix for separation and self-assembly is a scheme using polymers consisting of a carbon backbone because of their structural compatibility. The electric conductivity of the filler-based polymer composite is substantially increased with CNT loading. Raman frequency shifts have demonstrated their noncovalent interfacial interactions, through tensile and strain as well as compression transferred by the outer and inner layers of the multiwalled nanotube (MWNT), respectively.[2,3] These interfacial interactions, along with the high aspect ratio (D/L in the order of 1:1000), are believed to enhance conduction path through hopping and tunneling in the composite.

The study and design of the CNT/polymer composite emerged with increased knowledge of the interactions in the composites. Most recently, CNT-based composites appear promising for applications in telecommunication,[4] actuators,[5,6] photovoltaic cells,[7] light-emitting diodes (LEDs),[8] field emission[9] and radiation shielding.[10] The performance and lifetime of devices are enhanced through CNT composites. This also opens a venue for many nonconducting polymer applications with good structural stability and flexible processability. Because mobility can be increased to avoid electrostatic effect, the composites can be used for passive electronic components as well as electromagnetic shielding.[11] By forming the CNT composites, we overcome the disadvantages pertinent to nonconducting polymers to improve their performance as a result of the superior electric, thermal, and mechanical properties of CNT.

RAMAN SPECTROSCOPY BACKGROUND

The interfacial molecular structure studied by the high-throughput micro-Raman probes is correlated with macroscopic physical properties. In particular, resonant Raman spectroscopy probes the electronic density of states (DOS) singularity band gap by tuning the Raman excitation energies. The optical approach has the advantage of eliminating the contact between the electrode and the nanometer-scale material. It has become a very powerful technique to study the electronic properties of nanostructured materials.

Fig. 1 shows the energy level diagram of the Raman effect. When the monochromatic light excites matter (solid, liquid, or gas), it scatters light as a result of relaxation from virtual states, which are short-lived and distorted by the oscillating electric field of the incident light. A photon is immediately emitted or absorbed during the scattering. Depending on higher or lower vibration states, it relaxes back, resulting in stoke and antistoke scatterings, respectively. A small percentage of light scatters the same frequency as the incident light when it relaxes back to the same vibration state as the initial state (called Rayleigh scattering). When the incident light matches the electronic transition states, resonance scattering occurs with an intensity of up to 10^8 over the nonresonance intensity. Thus the electronic band gap can be measured.

Raman spectroscopy emerged as a powerful tool to study carbonaceous materials in recent years with the breakthrough in charge-coupled detectors and laser radiation sources. Both polymer and CNT have very high

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013849

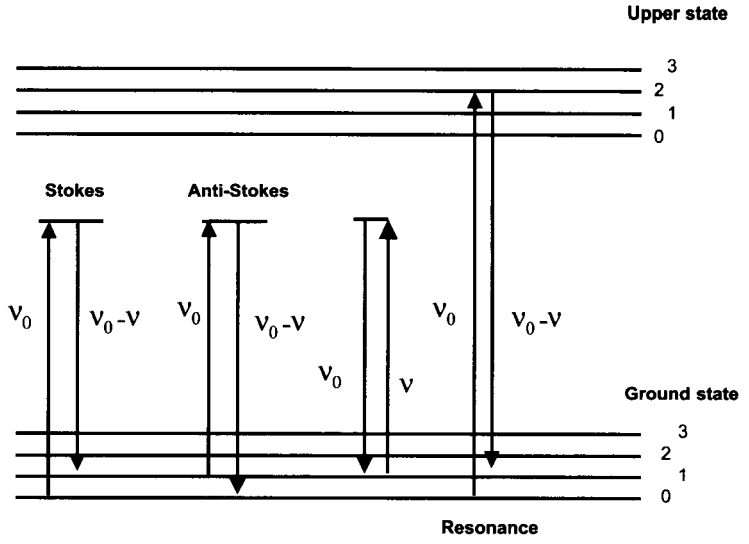


Fig. 1 Energy level diagram of Raman scattering effect.

percentages of the chemical composition of carbon. The high polarizability intensity of the carbon bond (C–C or C=C) results in especially strong Raman scattering in these classes of materials. Although infrared (IR) spectroscopy is frequently used for probing the polymer side-chain vibration symmetry, Raman spectroscopy is an ideal technique to study the properties of polymer backbones, the CNT lattice structure, and their interactions in composites.

Raman spectroscopy has been used to study the polymer vibration structure of backbone composition, conformation, and crystallinity. Targeting the conjugation length, chain conformation, interchain interaction, and polaron–bipolaron ratio, the correlation of the electronic and

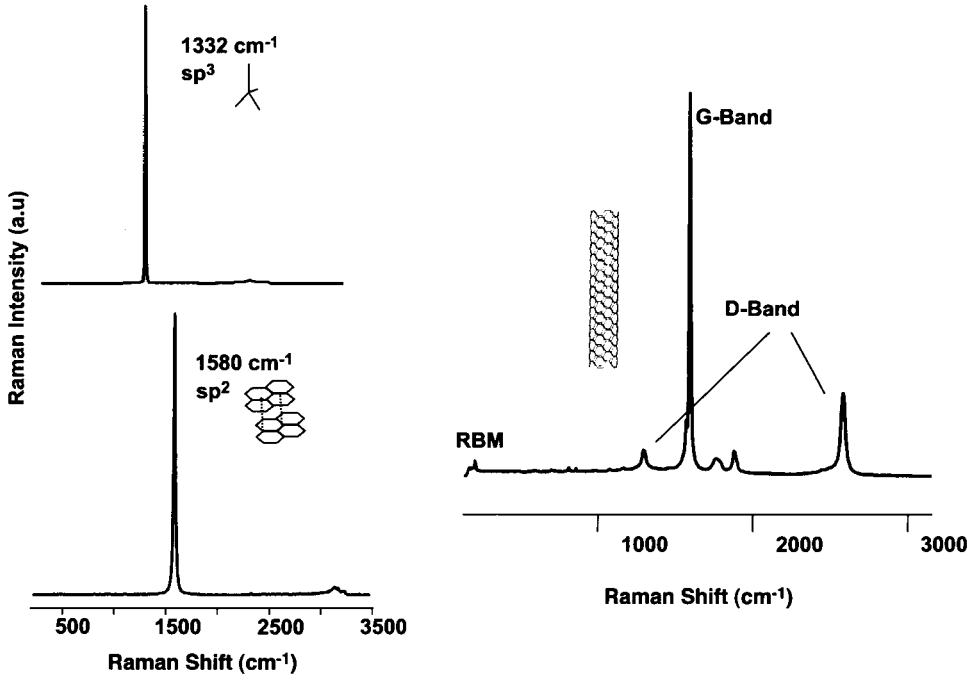


Fig. 2 Raman spectra of crystalline diamond, graphite, and SWNT in visible excitations.

chemical structural changes of polymers can be established.[12] In the CNT composite, the polymer structure changes caused by the CNT loading can be monitored by the Raman frequency shift and lines shape analysis.

Fig. 2 shows the Raman spectra of crystalline diamonds, graphites, and SWNTs. Like its parent forms of diamond and graphite, SWNT also processes distinct spectroscopic signatures. The characteristic D-band and G-band correspond to sp^3 and sp^2 vibration modes displayed in diamonds and graphites, respectively. The quasi one-dimensional (1-D) CNT also displays nanometer-scale signatures in tangential modes (TMs) and radial breathing modes (RBMs) regions. The rich information about structure provided by Raman spectroscopy has been also demonstrated in probes of CNT chirality, size, and defects.

A detailed analysis of SWNT structures and SWNT Raman modes can be found in an excellent review[1] and the entry on "Single-Walled Carbon Nanotubes: Geometries, Electronic Properties, and Actuation." Briefly, SWNT is a graphene sheet rolled into a cylindrical shape with very high aspect ratio. The axial symmetry exhibits the spiral conformation called chirality. The CNT belongs to D_{nh} (n is even number), D_{nd} (n is odd number), and C_n (achiral) groups. The G-band is formed through graphite Brillouin zone folding. According to group theory, the TM exhibits A_1, E_1, and E_2 symmetries that are Raman-active; the characteristic Raman signature consists of TMs including G-bands around 1590 cm^{-1} and D-bands around 1350 cm^{-1}; another signature band is the second-order Raman mode at 1730 cm^{-1} in the single-walled CNT. In the TM region, the line shape displays a symmetric profile around 1590 cm^{-1} for semiconductor SWNTs; a broad asymmetric G-band around 1560 cm^{-1}, as well as a sharp band around 1580 cm^{-1} are characteristic features associated with metallic SWNT sin the bundle.[13] Usually, Breit–Wigner–Fano (BWF) line shape is used to fit the lower-frequency component of the G-band. This phonon and electron continuum Raman line shapes can be obtained from:

$$I(\omega) = I_0\{1 + (\omega - \omega_0)/q\Gamma\} \times 2/\{1 + ((\omega - \omega_0)/\Gamma)_2\} \quad (1)$$

where I_0, ω_0, $1/q$, and Γ are intensity, renormalized frequency, the interaction of the phonon with continuum states, and the broadening parameter, respectively.[14]

The energies of allowed optical transitions (ATOs) are determined by both the diameter and the chirality of the SWNT,[15] as schematically illustrated in Fig. 3.

Below 500 cm^{-1}, the frequencies are contributed from the RBM Ω. The SWNT diameter d is computed[16,17]

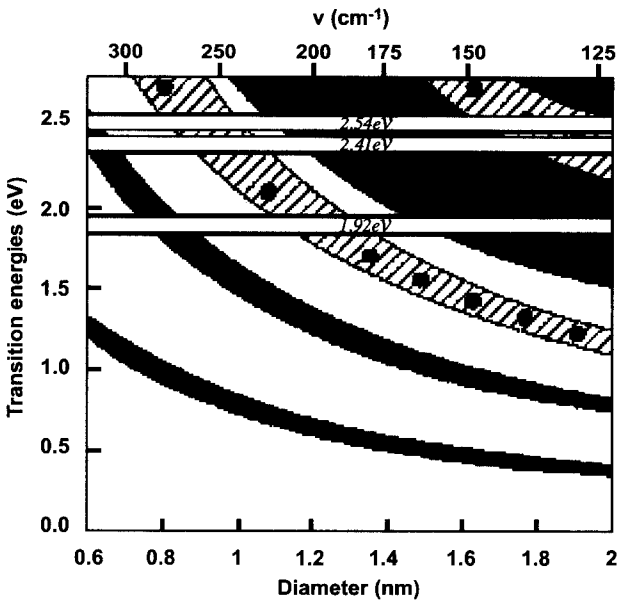


Fig. 3 Allowed optical transitions for SWNT with various diameters and chiralities calculated in the zone folding scheme. Black areas correspond to semiconducting tubes; dashed areas correspond to metallic tubes. The laser energies are labeled in write frames. The maximum of Raman intensity is expected for RBMs and TMs at the high and low limit of the frames, respectively. (©APS, 2001.)

from the low-frequency Raman spectra region using a phenomenological Eq. 2:

$$\Omega_{RBM}\ (cm^{-1}) = \alpha/d\ (nm), \quad \text{where } \alpha = 248\ cm^{-1}\ nm \quad (2)$$

The MWNT consists of tens of SWNTs. The stacking and assembly of each tube rarely align perfectly. Two forms are commonly observed in experiments: bamboo structure and cone shape stacking. Tube stacking generally lowers the symmetry of the CNT. The Raman mode resulting from these two stacking assemblies gives the secondary defect-induced mode at about 1612 cm^{-1} as a shoulder peak in the D-band. The cap at the end of the tube is usually open with the larger-diameter CNTs and chemical treatments. This feature can be monitored for MWNTs undergoing purification and dispersion because of their interactions with chemical agents and the chemical environment. The G-band to D-band ratio is substantially lower than that of SWNTs.

The confocal microscopic Raman probe enables the study of unique properties exhibited in nanometer-scale composite materials in the volume order of 1 μm. Various techniques such as polarized Raman and Raman

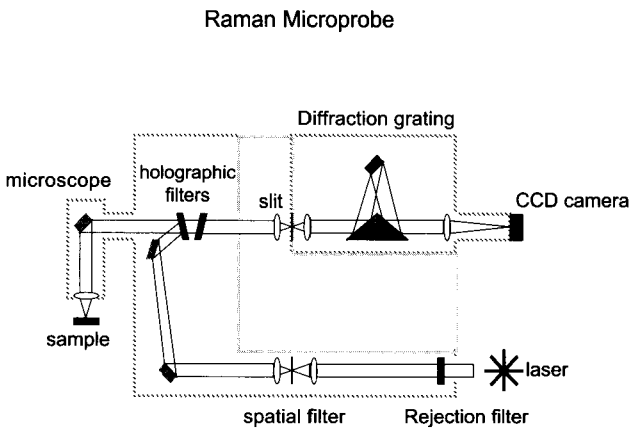


Fig. 4 The schematic illustration of a typical micro-Raman instrument setup. (Courtesy of Renishaw, Inc.)

imaging[18] became possible for the study of SWNT interactions in nonconjugated polymer matrices in high spatial resolution. The interfacial van der Waals interaction of SWNTs in the polymer composite demonstrates important contributions to many physical properties.

The Raman instrument is generally composed of three parts: the excitation light source, the spectrometer, and the detection apparatus. Fig. 4 is the schematic setup of a typical micro-Raman from Renishaw, Inc. The setup enables relative low laser power (2–3 mW) focused on the 1-μm^2 sample spot, providing the necessary high spatial resolution for nanomaterial studies.

SYNTHESIS OF CNT/POLYMER COMPOSITES

In Situ Growth

It is a challenge to control the growth process, but managing it effectively is nevertheless an effective way to obtain homogeneous composite films and to achieve a uniform CNT dispersion.

UV and chemical-polymerized composites were processed by polymerization of monomers in the presence of SWNTs (an initiator sometimes is added as needed) in the sonication bath. The same polymerization can incorporate SWNTs into a polymer matrix through ionic surfactant solution, with a concentration slightly higher than the critical micelle concentration. This second polymerization approach is so-called "emulsion-processed." Either oil or water-soluble initiators maybe added to accelerate the polymerization. The third method is electropolymerization. The CNTs will be used either as the electrode onto which the polymer is deposited, or as part of the electrolyte to be codeposited. When the CNT array is used as the electrode, the composite formed through the polymerization occurs on the surface of the CNT electrode. As the nanotube is dispersed with a monomer in the organic solution as a mixed electrolyte, the composite forms homogeneously onto a conventional electrode such as platinum. The redox potential, oxidation–reduction cycle time, and choice of dopant ions[19,20] in the electrolyte and the functional group in the CNT all act as control parameters to adjust composite film thickness and conductivity. Conducting polymers, such as polypyrrole (PPy), are commonly synthesized with either scheme.[21,22]

Shear Field Processing

SWNTs are added to a solution of polymer in toluene or *N,N*-dimethylformamide solution. High sonic power is briefly (a few minutes) used to break up the CNT aggregates, and then low power is applied to the sonic bath for a few hours. Generally, a few percent of CNTs remain in the solution other than those precipitated as aggregates. After a few days of settlement, the polymer/CNT composite from the supernatant can either be piped out and cast as thin film, or further thermally blended. In the latter, the composite can be dried in the oven and then melt-pressed above its glass transition temperatures a few times to ensure uniform mixing. The anisotropic effect introduced in the melt-blended process may seem predominant than in situ-polymerized composites. Some researchers[23] believe that the different thermal expansions between SWNTs and PMMAs can significantly contribute to their interfacial adhesion. This causes stress and improves alignment between the two components in the melt-blended composite in anisotropic directions. For example, molecular dynamic simulations of polyethylenes and SWNT composites show that the polyethylene diffusion coefficient is 30% higher in the direction in which its backbone is parallel to the SWNT axial direction than perpendicular to it.

INTERFACIAL INTERACTIONS IN THE COMPOSITE OF INCREASED ELECTRIC CONDUCTIVITY

Some of the processing conditions have subtle, yet profound, impacts on composite properties. Raman studies show that applied sonication shear field strength during CNT composite formation can introduce defects to the CNT sidewall, change its diameter,[24,25] or even open the end caps on the tubes. The defect sites are usually the most

active sites where the charge transfer most likely occur. The chemical interaction usually occurs on the CNT tip because the fullerene-like cap is more reactive. The less invasive noncovalent interaction generally happens along the CNT sidewall in the form of van der Waals, hydrogen bond, or electrostatic interactions, all of which are more commonly observed in the polymer composites. Raman analysis is a very sensitive technique for investigating these interactions. The G-band and D-band ratio in the Raman shift can illustrate processing and growth-related mechanisms. However, the backbone conformation changes of the polymer in conjunction with the diameter changes of the CNT routinely probed by Raman spectroscopy are likely the causes of the constituent components electronic conductivity increases in the polymer composites.

In many cases, the measured conductivity change depends on interactions on the molecular level for nanometer-scale composites. Raman spectroscopy has demonstrated that the electronic DOS band gap of CNTs in the composite can be influenced by interactions formed during the composite formation.

PMMA/SWNT Composites

The polymethyl methacrylate (PMMA/CNT) composites are among the most studied composites. The electronic conductivity of the MWNT composite is increased by several orders of magnitude.[26] Because there is no conjugated bond in PMMA to interact with the CNT, the interfacial interactions are likely to play important roles in composite formation. Therefore this is an ideal system to investigate the structure and conductivity changes simultaneously. These changes can be interplayed with direct conductivity measurements, as well as optical probes to electronic DOS through resonant Raman spectroscopy.

The PMMA/SWNT composite is fabricated into a spun fiber by the melt-processed film and analyzed by Raman spectroscopy.[27] The SWNT could either position itself to the fiber axis, or isotropically disperse in the composite. Fig. 5 shows the polarized Raman spectra of the composite. The intensity dependence of the RBMs with respect to the laser incident angles demonstrates that the SWNT is highly anisotropically aligned with the fiber axis. This corresponds to the much higher conductivity directly measured in the fiber axial direction, as opposed to the lower conductivity perpendicular to the processing flow direction.

Apparently, further investigations are needed of the PMMA and SWNT structural details that affect the change of conductivity. The changes of SWNT sizes with loading have been studied with Raman spectroscopy by using multiple wavelength excitations.[28] At both 514- and 1064-nm excitations, Fig. 6a shows that SWNT RBM

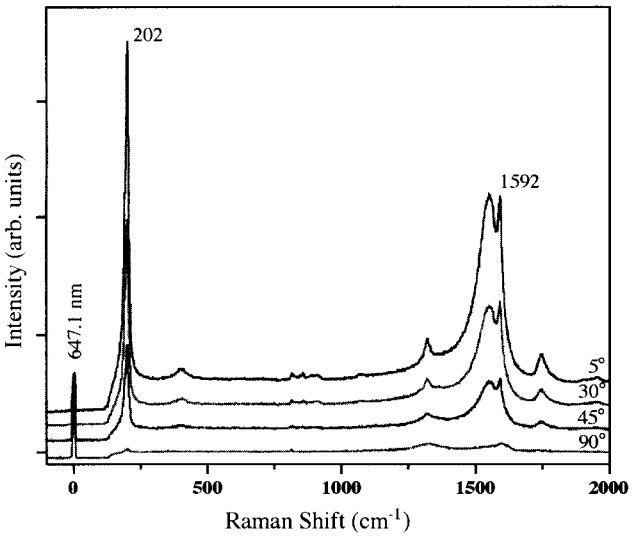


Fig. 5 Raman spectra of SWNT/PMMA nanocomposite fiber at the indicated fiber angles with respect to the incident laser polarization axis. Spectra were recoded using the 647.1-nm excitation, with the analyzed polarization parallel to the incident polarization. The angle-dependent intensity of the breathing mode of the SWNTs (202 cm^{-1}) was used to determine the extent of nanotube alignment relative to the fiber axis. (©Elsevier Science B.V., 2000.) (*View this art in color at www.dekker.com.*)

frequency upshifts toward larger-diameters distributions as calculated from Eq. 2. The PMMA composites with SWNT loadings of 2.5%, 5%, 10%, and 20% are also studied with resonant conditions at 1064 and 647 nm for semiconducting and metallic SWNTs, respectively. These two excitations correspond to the first semiconducting and metallic DOS singularity bandgaps, respectively. In the 1064-nm excitation that corresponds to the semiconducting SWNT, Fig. 6b shows that among the five components in the TM region (1553, 1568, 1573, 1592, and 1597 cm^{-1}), the intensities of the three peaks located at 1553, 1568, and 1573 cm^{-1} decrease in the composites. More drastically, in the 633-nm laser excitation that corresponds to the metallic SWNT, the relative ratio of 1543/1558 cm^{-1} in the G-band is reversed from the pristine SWNT. This is an indication of the increased percentage of metallic SWNT because the G-band line shape featured by the BWF fit in Eq. 1 in the metallic SWNT resonant condition. The other possibility is that PMMA/SWNT interaction replaces the interface in the SWNT bundles. Furthermore, the interaction between SWNT and PMMA is investigated in full width at half maximum (FWHM). The FWHM increases with SWNT loadings, and decreases with excitation wavelengths go from 458, 514, 676, to 1064 nm. The observed G-band blue shifts and the SWNT diameter increases, clearly

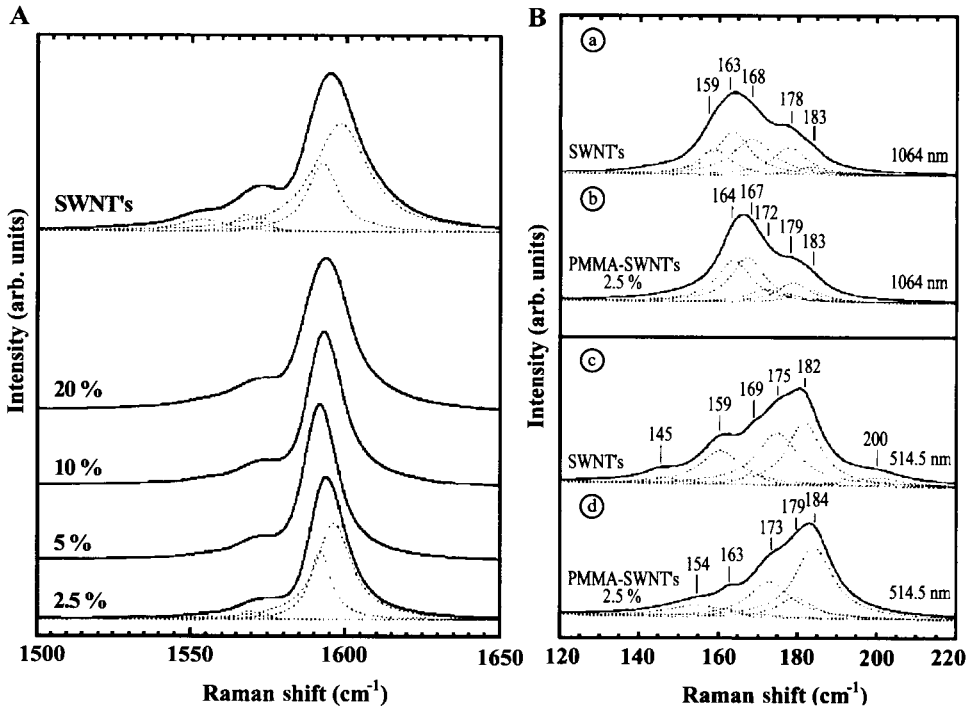


Fig. 6 (A) Raman spectra of PMMA/SWNT composites and SWNT films in the frequency range 1500–1650 cm^{-1} with 1064 nm excitation. The spectra are resolved for PMMA/SWNT composites of 2.5% concentration and SWNTs. (B) Decomposition of low-frequency bands: (a) SWNT film with 1064 nm excitation; (b) PMMA/SWNT film for 2.5% concentration with 1064 nm excitation; (c) SWNT film with 514.5 nm excitation; and (d) PMMA/SWNT film for 2.5% concentration with 514.5 nm excitation. (©Elsevier Science S.A., 2000.)

demonstrating the van der Waals interaction between SWNT/PMMA in the composites, in place of intertube interactions in the pristine SWNT.

The interfacial interaction in SWNT/PMMA composites can be further illustrated by Raman studies on two differently processed composites.[29] Fig. 7a gives the Raman results of melt-processed and UV-polymerized composites in the metallic resonant excitation energy at 633 nm (1.96 eV). SWNT Raman spectra are used for reference. Comparing to both UV-polymerized composites and pristine SWNTs, the melt-blended composite exhibits a significant difference in the SWNT G-band line shape profiles around 1590 cm^{-1}. The G-band in UV-polymerized composites retains the same Lorentzian line shape as that of the pristine SWNT. The broader G-band feature in the thermally treated composite can be fit with the BWF line shape for the SWNT diameter between 1.2 and 1.6 nm in the AOT,[30] as discussed in Fig. 3. The BWF line shape is the signature of metallic SWNTs resulting from the coupling between the phonon and the electronic continuum.[31] The predominant metallic feature shown in Fig. 7a demonstrates a significant increase of the metallic SWNT percentage in the melt-blended composites. The SWNT in the UV-polymerized composite, in which PMMA exhibits no anisotropic effect, retains the same chirality as in the pristine SWNT sample.

The anisotropic interfacial interaction between PMMA and SWNT in the melt-blended composite can be monitored by the PMMA backbone conformation change, especially when the C–H stretching in the backbone is amplified by Fermi resonance.[32,33] The stretching symmetry contributions from the backbone relative to the side chain are significantly different in the two composites examined. The relative ratio of the $C–H_2$ stretch bands to the side-chain C–C and C–O stretching mode is 30% higher in the melt-blended composite than that in the UV-polymerized composite. Schematically, the $C–H_2$ stretching is little affected by the anisotropic SWNT/PMMA interactions in Fig. 7b; however, it is hindered by isotropic SWNT/PMMA interactions in Fig. 7c. Because the side-chain vibrations are not sensitive to the orientation of PMMA/SWNT interaction, the higher intensity of the $C–H_2$ stretching band suggests the likely stronger anisotropic interaction between PMMAs and SWNTs. This interaction is believed to cause metallic SWNT enrichment in the melt-blended composite.

The electric conductivity increase investigated by Raman spectroscopy was also confirmed by composite

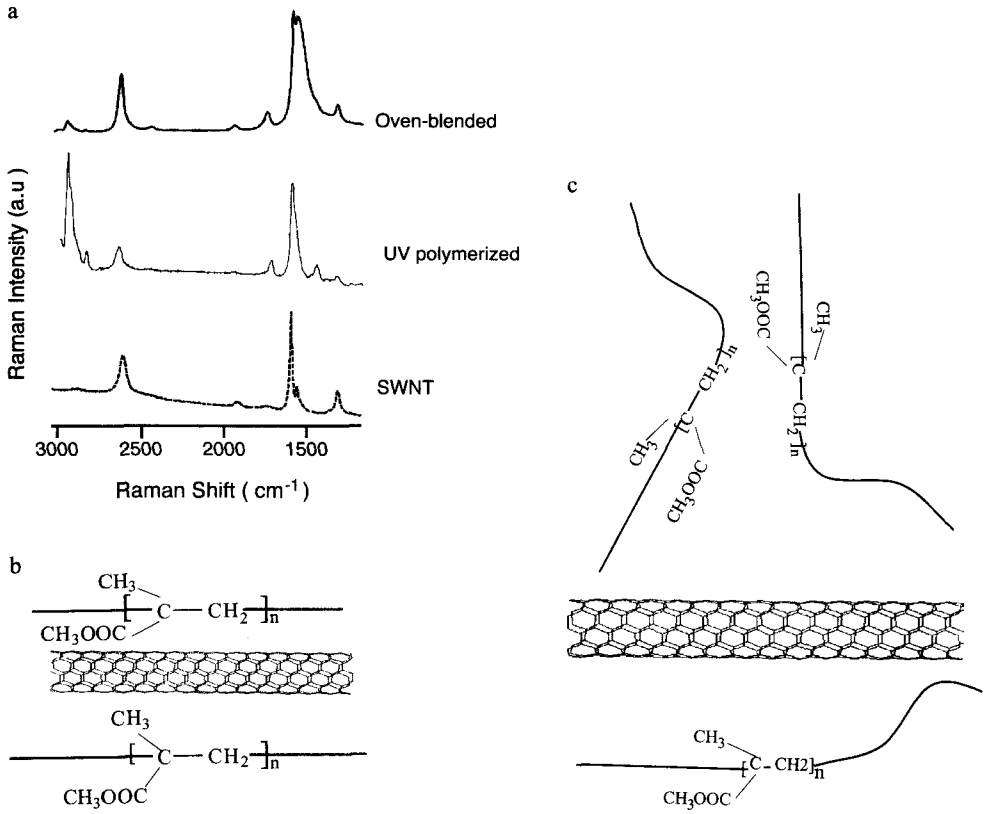


Fig. 7 (a) Raman spectra of the two PMMA/SWNT composites and pristine SWNT excited at 633 nm. The G-band of SWNTs in the melt-blended composite around 1590 cm^{-1} is distinctly different. This line shape fits with the BWF line shape and is recognized as the signature for a metallic SWNT. The G-band in UV-polymerized composites retains the same Lorentzian line shape as in pristine SWNT. (b, c) Schematic illustrations of the PMMA and SWNT anisotropic and isotropic interactions in the melted-blended and UV-polymerized composites, respectively. The anisotropic interaction between SWNT side wall and PMMA backbone has little effect on side-chain vibrations, but will affect the backbone C–H_2 stretch because axial stretching can be blocked by SWNT/PMMA if not aligned.

dielectric measurements.[34] The higher dielectric constant is obtained for the melt-blended composite than for the UV-polymerized composite. This is consistent with the possible dipole interaction enhanced by the anisotropic interaction between the PMMA and SWNT in the melt-blended composite. The strain-induced band gap change[35] possibly could also contribute to the increase of the metallic percentage of SWNTs in the melt-blended composites.

Other Polymer Composites

The similar interfacial interaction effect on the conductivity in PMMA/SWNT composites has been observed in several other polymer composites.

The in situ-polymerized rigid polymer poly(*p*-phenylene)benzobisoxazole (PBO)/SWNT composite is similarly prepared as UV-polymerized PMMA/SWNT composites.[36] Its measured conductivity remained the same as that of the neat polymer, possibly because of a lack of strong interfacial interactions in the composite, which is not detectable from Raman studies.

The conductivity of polystyrene and SWNT composites formed from a miniemulsion process increases by ten order of magnitude at 8.5% SWNT loading.[37] The interactions between polystyrene and SWNT in the composites are both exhibited in the Raman spectra. The SWNT G-line at 1580 cm^{-1} shifted to higher frequencies whereas the RBM upshifted. The loss of intensity of polymer C=C stretch modes at 1670 cm^{-1} is also observed among the polystyrene Raman modes.

Polyimide (PI)-based composites appeared to increase electric conductivity with SWNT loading. Conductivity percolation is reached with SWNT loading between 0.02% and 0.1%.[38] The interactions in CNT/PI nanocomposites can be controlled to increase polaron mobility and quantum efficiency. The Raman results of this

photovoltaic electrolyte material suggest the charge transfer mechanism correspondingly.[39]

CHARGE TRANSFER AND STRONG π INTERACTION IN COMPOSITE WITH INCREASED ELECTRIC CONDUCTIVITY

The SWNT G-band is very sensitive to the doping and charge process. The Raman frequency shifts depending on the types of donor or acceptor. The charge transfer between halogen or alkali metal and SWNT was first demonstrated by Raman studies[40] shown in Fig. 8. The TM frequency shifts in the SWNT sidewall around 1550 cm^{-1} show the changes in accordance with the corresponding diameter changes in RBM. The charge transfer soften mode is identified as BWF interference in the alkali metal-doped SWNT bundle. The π electron

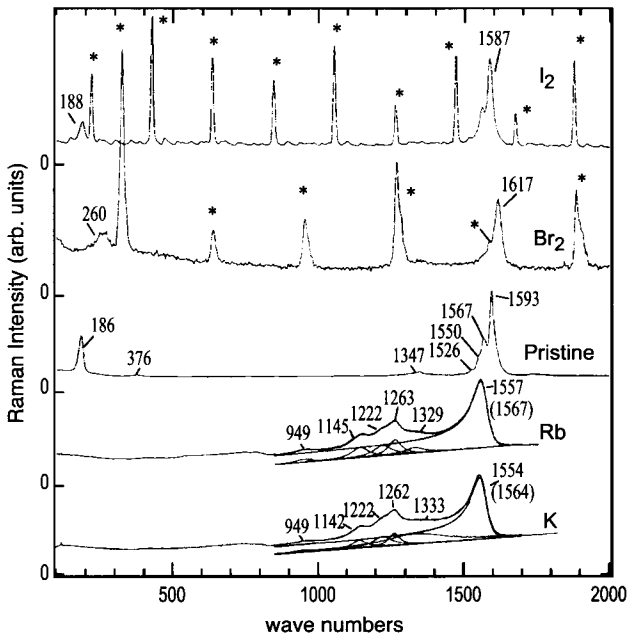


Fig. 8 Raman scattering spectra for pristine SWNT bundles reacted with various donor and acceptor reagents. From top to bottom: I_2Br_2, pristine SWNT, Rb, and K. The backscattering spectra were taken at T = 300 K using 514.5 nm radiation. In the spectra for both halogen-doped SWNT bundles, a harmonic series of peaks (indicated with asterisk) is observed, which are identified with the fundamental stretching frequency ω_s: ~220 cm^{-1} (I_2) and ~324 cm^{-1} (Br_2). The spectra have been scaled so that the strongest SWNT feature appears to have the same intensity. In the vicinity of the strongest high-frequency mode around 1550 cm^{-1}, the Raman spectra for SWNTs doped with K or Rb are fitted with superposition of Lorentzian functions and an asymmetric BWF line shape on a linear continuum. (©Macmillan Publishers, 1997.)

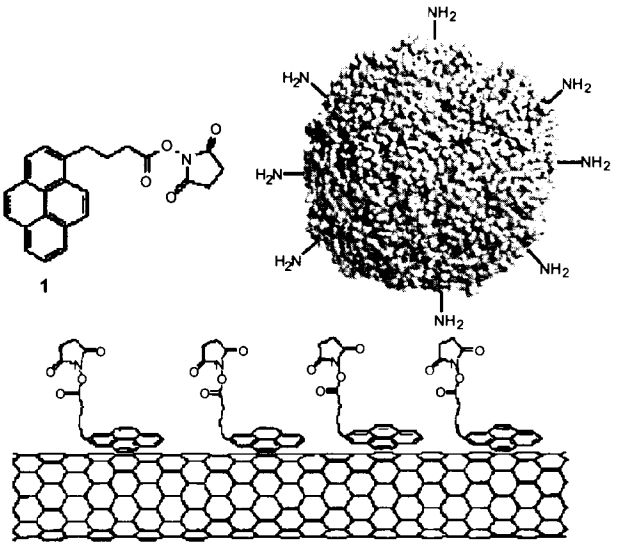


Fig. 9 1-Pyrenebutanoic acid succinimidyl ester 1 irreversibly adsorbing onto the sidewall of a SWNT via π-stacking. (©ACS, 2001.)

transfer from the carbon to the acceptor (Br_2, I_2) or donor (K, Rb) is believed to introduce both TM and RBM changes.

The π-stacking of the planar pyrenylene π-moiety in the 1-pyrenebutanoic acid succinimidyl ester with the CNT graphite basal plane forms a very stable self-assembled composite,[41] which changes the SWNT from hydrophobic to hydrophilic nature. This structure interaction was schematically illustrated in Fig. 9. Although Raman studies are not reported for this composite, the strong interaction was investigated by another similar polymer composite in terms of charge transfer. The Raman studies of composites of the aromatic perylene-doped SWNTs are believed to have a charge transfer in the G-band modes.[42] Fig. 10 shows the similar Raman shifts as observed in alkali and halogen doping in the SWNT. Although the strong perylene modes, such as 1296 cm^{-1}, are not observed in the areas investigated by Raman spectra, the inactive SWNT vibrational modes around 1542, 1549, and 1610 cm^{-1} in the G-band emerge in the composite and exhibit different degrees of enhancement because of perylene doping interaction.

The PmPV composite is well studied by various spectroscopy techniques. PmPV is an important photovoltaic and LED material in optoelectronic applications. The CNT composite is expected to improve device performance by optimizing the SWNT loading, tube length, and its chirality. The scanning tunneling microscopy (STM) results indicate that PmPV, as a coiling polymer, is interacting with SWNT with periodical surface wrapping,[43] as shown in Fig. 11. The wrappings

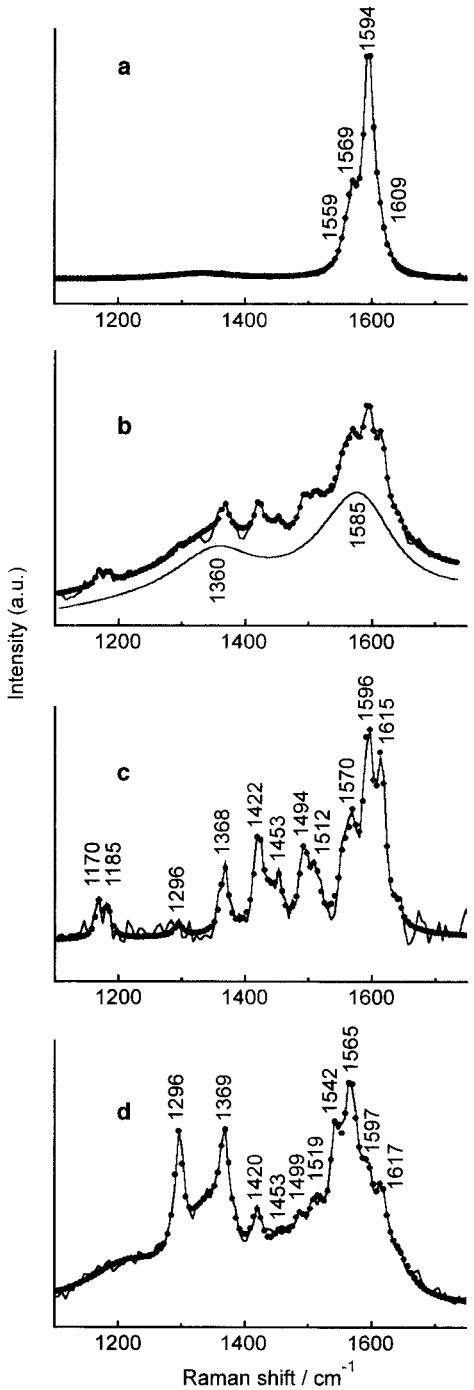


Fig. 10 Typical Raman spectra in different areas. (a) Spectrum of pristine nanotube G-band is fitted by Lorentzian curves peaking at 1559 cm^{-1} (E_{2g} mode), 1569 cm^{-1} (A_{1g}), 1594 cm^{-1} (A_{1g}+E_{1g} modes), and 1609 cm^{-1} (E_{2g} mode). (b) Spectrum of perylene-doped tubes in zone B. (c) Signatures of perylene–tube interaction after subtraction of the asymmetric background and of the growing fluorescence background in zone B. (d) Spectrum of perylene-doped tubes in another area. (©Elsevier Science B.V., 2002.)

are refractive and conform to the sidewall lattice structure of the SWNT. The electric conductivity increased up to eight orders of magnitude at 8.5% MWNT loading. At less than 7% SWNT loading, an increase of one order of magnitude to 1.5×10^{-9} S/m was observed; between 7% and 11% loading, an increase in conductivity of eight orders of magnitude has been observed; whereas the conductivity plateaus to 3 S/m after 36% loading.[44] The interactions in the composite possibly enhance polaron mobility through the change of the polymer backbone conformation and through a better hopping and tunneling path. It is believed that PmPV significantly reduces electron delocalization and vibrational freedom after the interaction with SWNT in the composite.

The dependence of increased conductivity on CNT loading in the composites is established by Raman studies under resonant conditions of the semiconducting and metallic SWNTs, respectively.[45] At 676 nm wavelength excitation resonant with the metallic SWNT, the polymer vinylene stretch mode at 1627 cm^{-1} evolves into two well-defined modes at 1630 and 1625 cm^{-1}. The SWNT TM of 1598 cm^{-1} is simultaneously downshifted to 1593 cm^{-1}, whereas the 1568 cm^{-1} mode disappears. The RBMs measured with laser excitation energies resonant at the semiconducting condition display the upshifts by about 7 cm^{-1}; the relative intensity ratio of the individual frequencies is also changed. The

Fig. 11 STM image of two PmPV-coated SWNTs lying on HOPG substrate. The ordered coating can be seen in the tube on the right. There are two individual tubes on the left, separating at the bottom. (©ACS, 2002.)

SWNT diameter distributions derived from Eq. 1 indicate that the diameters shift toward smaller size distribution in the composites. However, the D-band to G-band ratio is significantly changed, which suggests that sp^3 carbon may act as the active site that interacts with PmPV. The spectroscopy evidence of these vibration symmetry changes from both PmPV and SWNT, combined with STM studies, show the helix wrapping of PmPV backbone around the MWNT sidewall with strong interactions. The luminescence study shows no quenching, which is usually associated with the charge transfer mechanism. Therefore it can be concluded that the composite conductivity increases are because of the PmPV backbone's strong π–π interactions with the SWNT sidewall, rather than charge transfer mechanism.

The majority of the pristine LED polymer materials needed to be doped to achieve high-enough carrier mobility for device application. The polaron and bipolaron ratio not only reflects the conductivity of the composite, but also is the direct result of the symmetry change of the polymer because of the interaction of the CNT with the polymer.

The electropolymerized polypyrrole (PPy)/CNT thin film composite has attracted considerable attention. Being a conjugated polymer, PPy has strong electron and phonon interactions (called polaron) in the lattice. It resembles an ionic lattice, so bipolaron can form through interactions between neighboring polarons. Two significant changes in the MWNT symmetry mode are shown in the curve-fitting results of Raman spectra at 633 nm excitation.[46] First, the defect-induced MWNT shoulder peak D′ at 1612 cm^{-1} substantially decreased after the composite film formation. The normalized intensity ratio of D-band to G-band significantly decreases from 2.5 in the MWNT arrays to 0.4 in the composite film. This indicates that the MWNT sidewall sp^3 graphitic carbon interacts with PPy. The observed decrease of the D-band intensity is similarly demonstrated in the PmPV/SWNT composite as well.[8] Therefore the symmetry of both MWNT and PPy is expected to change. Examine the bipolaron (labeled with *) and polaron (labeled with †) pair[47] in Fig. 12. MWNT composites have much higher bipolaron contributions than in the CPE composite. Furthermore, bipolaron-only[48] mode at 1247 cm^{-1} is only observed in the MWNT/PPy composite film. However, the polaron mode at 1044 cm^{-1} is more predominant in the CPE/MWNT film. In addition, the skeletal band[49] C–N stretching mode at 1487 cm^{-1} shoulder is higher in the MWNT/PPy composite, indicating stronger backbone vibrations. All the Raman results suggest better PPy backbone conjugation in the MWNT composite than that in the CPE composite. When PPy electropolymerized onto the highly orientated electrode array, its backbone conforms to the MWNT sidewall. Therefore the polymer chain deformations,

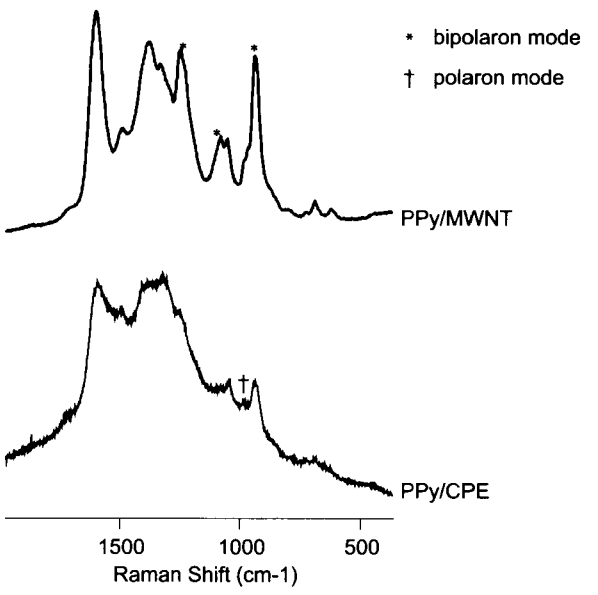


Fig. 12 Raman spectra of PPY electropolymerized at (a) MWNT array electrode, and (b) at carbon paste electrode. In the MWNT/PPy composite, the dipolaron modes at 986 and 1051 cm^{-1} (labeled with *) corresponding to polaron modes (labeled with †) 986 and 1051 cm^{-1} are much higher than that in the CPE/PPy composite. The bipolaron mode at 1247 cm^{-1} is only observed in the MWNT/PPy composite. However, the polaron mode at 1044 cm^{-1} is more predominant in the CPE/MWNT film. The skeletal band C–N stretching mode at 1487 cm^{-1} shoulder is higher in the MWNT/PPy composite, indicating stronger backbone vibrations.

kinks, twists, and angle distortions between aromatic ring planes can be reduced. The high bipolaron density enhances conjugation in the MWNT/PPy composite film. Thus the resultant film is structurally very stable and highly conductive. This preferred configuration is confirmed by the SEM image of the MWNT after composite formation. The remarkable uniformity of this widely studied[50,51] composite film can be understood because of its unique nanostructural interactions.

Recently, Raman analyses have become a powerful technique to identify the separation of metallic and semiconductor SWNTs in both the bulk quantity and the nanoscale approach.

Resonant Raman analysis is used to probe the conductivity of SWNTs treated in the polar organic solvent, with octadecylamine (ODA) possibly self-assembled onto SWNTs.[52] The Raman line shape analysis around frequency of TMs is used for gauging semiconducting and metallic features under the respective resonant conditions. The resonant Raman spectroscopy concludes the predominant semiconducting and metallic compositions in the supernatant and precipitant, respectively.

Raman study of CNT conductivity has also been successful in nanodevices.[53] Both RBM and TM frequencies are used to analyze compositions in the electrode and in the suspension after the a.c. electric field on a pair of microelectrodes laid out on the oxidized Si substrate. The induced dipole moment migrates the metallic SWNTs to the electrodes, leaving the semiconducting SWNT in the suspension. The polarized Raman also shows the highly aligned nanotube along the electric field. In this application, the in situ microscopic Raman probe demonstrates the ideal technique to study CNT conductivity beyond macroscopic chemical environments.

CONCLUSION

Recent research progress in Raman spectroscopy demonstrates that it is an ideal technique to study the electronic conductivity of the CNT and polymer composites in two approaches. In the first approach, the vibration symmetry studies of the CNT and polymer are directly correlated with conductivity measurements, providing possible mechanisms of the composite electronic conductivity change caused by nanostructural interactions. In the second approach, Raman spectroscopy directly probes the composite electronic DOS band gap through resonant scattering. The implications of interfacial interaction, as well as charge transfer between the constituent components in the composite are studied by Raman spectroscopy. The Raman spectroscopy study will guide processing conditions, including thermal treatment and shear field strength, as well as polymerization specifications for better control in electronic and optical applications.

REFERENCES

1. Saito, R.; Dresselhaus, G.; Dresselhaus, M.S. *Physical Properties of Carbon Nanotubes*; Imperial College Press: London, 1998.
2. Wagner, H.D.; Lourie, O.; Feldman, Y.; Tenne, R. Stress-induced fragmentation of multiwall carbon nanotubes in a polymer matrix. Appl. Phys. Lett. **1998**, *72*, 188.
3. Schadler, L.S.; Giannaris, S.C.; Ajayan, P.M. Load transfer in carbon nanotube epoxy composites. Appl. Phys. Lett. **1998**, *73*, 3842.
4. Chen, Y.-C.; Raravikar, N.R.; Schadler, L.S.; Ajayan, P.M.; Zhao, Y.P.; Lu, T.-M.; Wang, G.-C.; Zhang, X.-C. Ultrafast optical switching properties of single-wall carbon nanotube polymer composites at 1.55 μm. Appl. Phys. Lett. **2002**, *81* (6), 975.
5. Landi, B.J.; Raffaelle, R.P.; Heben, M.J.; Alleman, J.L.; Van Derveer, W.; Gennett, T. Single wall carbon nanotube–nafion composite actuators. Nano Lett. **2002**, *2* (2), 1329.
6. Baughman, R.; Cui, C.; Zakhidov, A.A.; Iqbal, Z.B.J.N.; Spinks, G.M.; Wallace, G.G.; Mazzoldi, A.; De Rossi, D.; Rinzler, A.G.; Jaschinski, O.; Roth, S.; Keresz, M. Carbon nanotube actuators. Science **1999**, *284*, 1340.
7. Kymakis, E.; Amaratunga, G.A.J. Single-wall carbon nanotube/conjugated polymer photovoltaic devices. Appl. Phys. Lett. **2002**, *80* (1), 112.
8. Curran, S.A.; Ajayan, P.M.; Blau, W.J.; Carroll, D.L.; Coleman, J.N.; Dalton, A.B.; Davey, A.P.; Drury, A.; McCarthy, B.; Maier, S.; Strevens, A. A composite from poly(*m*-phenylenevinylene-*co*-2,5-dioctoxy-*p*-phenylenevinylene) and carbon nanotubes: A novel material for molecular optoelectronics. Adv. Mater. **1998**, *10*, 1091.
9. Alexandrou, I.; Kymakis, E.; Amaratunga, G.A.J. Polymer–nanotube composites: Burying nanotubes improves their field emission properties. Appl. Phys. Lett. **2002**, *80*, 1435.
10. Muisener, P.A. O'Rourke; Clayton, L.; D'Angelo, J.; Harmon, J.P.; Sikder, A.K.; Kumar, A.; Cassell, A.M.; Meyyappan, M. Effects of gamma radiation on poly(methyl methacrylate)/single-wall nanotube composites. J. Mater. Res. **2002**, *17*, 2507.
11. Chung, D.D.L. Electromagnetic interference shielding effectiveness of carbon materials. Carbon **2001**, *39*, 279.
12. Yacoby, Y.; Enrenfreund, E. Light Scattering in Solids VI. In *Topics in Applied Physics*; Cardona, M., Güntherodt, G., Eds.; Springer: Heidelberg, 1991; Vol. 68.
13. Rao, R.; Richter, E.; Bandow, S.; Chase, B.; Eklund, P.C.; Williams, K.A.; Fang, S.; Subbaswamy, K.R.; Menon, M.; Thess, A.; Smalley, R.E.; Dresselhaus, G.; Dresselhaus, M.S. Diameter-selective Raman scattering from vibrational modes in carbon nanotubes. Science **1997**, *275*, 187.
14. Klein, M.V. *Light Scattering in Solids I*; Cardona, M., Ed.; Springer-Verlag: Berlin, 1983; 169.
15. Brown, S.D.; Corrio, P.; Marucci, A.; Dresselhaus, M.S.; Pimenta, M.A.; Kneipp, K. Anti-stokes Raman spectra of single-walled carbon nanotubes. Phys. Rev., B. **2000**, *61*, R5137.
16. Jorio, A.; Saito, R.; Hafner, J.H.; Lieber, C.M.; Hunter, M.; McClure, T.; Dresselhaus, G.; Dresselhaus, M.S. Structural (*n*, *m*) determination of isolated single-wall carbon nanotubes by resonant Raman scattering. Phys. Rev. Lett. **2001**, *86* (6), 1118.
17. Alvares, L.; Righi, A.; Guillard, T.; Rols, S.; Anglaret, E.; Laplaze, D.; Sauvajol, J.L. Resonant

Raman study of the structure and electronic properties of single-wall carbon nanotubes. Chem. Phys. Lett. **2000**, *316*, 186.
18. Debarre, A.; Jaffiol, R.; Richard, A.; Tchénio, P. Raman hyperspectral imaging applied to chemical co-localization in diluted samples of perylene-doped nanotubes. Chem. Phys. Lett. **2002**, *366*, 274.
19. Liu, Y.; Hwang, B.-J. Identification of oxidized polypyrrole on Raman spectrum. Synth. Met. **2000**, *113*, 203.
20. Iriyama, Y.; Hanawa, M. Plasma polymerization of pyrrole and structures and properties of the polymerized films. Polym. J. **2001**, *33*, 419.
21. Chen, G.Z.; Shaffer, M.S.P.; Coleby, D.; Dixon, G.; Zhou, W.; Fray, D.J.; Windle, A.H. Carbon nanotube and polypyrrole composites: Coating and doping. Adv. Mater. **2000**, *12* (7), 522.
22. Chen, J.H.; Huang, Z.P.; Wang, D.Z.; Yang, S.X.; Wen, J.G.; Ren, Z.F. Electrochemical synthesis of polypyrrole/carbon nanotube nanoscale composites using well-aligned carbon nanotube arrays. Appl. Phys., A **2001**, *73*, 129.
23. Wei, C.; Srivastava, D.; Cho, K. Thermal expansion and diffusion coefficients of carbon nanotube–polymer composites. Nano Lett. **2002**, *2* (6), 647.
24. Cinke, M.; Li, J.; Chen, B.; Cassell, A.; Delzeit, L.; Han, J.; Meyyappan, M. Pore structure of raw and purified HiPco single-walled carbon nanotubes. Chem. Phys. Lett. **2002**, *365*, 69.
25. Liao, K.; Li, S. Interfacial characteristics of a carbon nanotube–polystyrene composite system. Appl. Phys. Lett. **2001**, *79*, 4225.
26. Sandler, J.; Shaffer, M.S.P.; Prasse, T.; Bauhofer, W.; Schulte, K.; Windle, A.H. Development of a dispersion process for carbon nanotubes in an epoxy matrix and the resulting electrical properties. Polymer **1999**, *40*, 5967.
27. Haggenmuller, R.; Gommans, H.H.; Rinzler, A.G.; Fischer, J.E.; Winey, K.I. Aligned single-wall carbon nanotubes in composites by melt processing methods. Chem. Phys. Lett. **2000**, *330*, 219.
28. Stephan, C.; Nguyen, T.P.; Lamy de la Chapelle, M.; Lefrant, S.; Journet, C.; Bernier, P. Characterization of single-walled carbon nanotubes–PMMA composites. Synth. Met. **2000**, *108*, 139.
29. Chen, B.; Cinke, M.; Li, J.; Meyyappan, M.; Chi, Z.; Harmon, J.P.; Muisener, P.A.O.; Clayton, L.; D'Angelo, J. Modifying electronic property of the single-walled carbon nanotube through anisotropic polymer interaction: A Raman study. Adv. Funct. Mater., submitted for publication.
30. Kataura, H.; Kumazawa, Y.; Maniwa, Y.; Umezu, U.; Suzuki, S.; Ohtsuka, Y.; Achiba, Y. Optical properties of single-wall carbon nanotubes. Synth. Met. **1999**, *103*, 2555.
31. Brown, S.D.M.; Jorio, A.; Corio. P.; Dresselhaus, M.S.; Dresselhaus, G.; Saito, R.; Kneipp, K. Origin of the Breit–Wigner–Fano lineshape of the tangential G-band feature of metallic carbon nanotubes. Phys. Rev., B **2001**, *63*, 155414.
32. Bulkin, B.J.; Krishnamachan, A. Vibrational spectra of liquid crystals: IV. Infrared and Raman spectra of phospholipid–water mixtures. J. Am. Chem. Soc. **1972**, *94*, 1109.
33. Gabor, B.P.; Peticolas, W.L. On the quantitative interpretation of bromembranal structure. Raman spectroscopy. Biochim. Biophys. Acta **1977**, *465*, 260.
34. Chen, B.; Li, J.; Lu, Y.; Cinke, M.; Au, D.; Partridge, H.; Meyyappan, M.; Harmon, J.P.; Muisener, P.A.O.; Clayton, L.; D'Angelo, J. Characterizations of enriched metallic single-walled carbon nanotubes in polymer composite. Macromol. Symp., submitted for publication.
35. Yang, L.; Han, J. Electronic structure of deformed carbon nanotubes. Phys. Rev. Lett. **2000**, *85* (1), 154.
36. Kumar, S.; Dang, T.D.; Arnold, F.E.; Bhattacharyya, A.R.; Min, B.G.; Zhang, X.; Varia, R.A.; Park, C.; Adams, W.W.; Hauge, R.H.; Smalley, R.E.; Ramesh, S.; Willis, P.A. Synthesis, structure, and properties of PBO/SWNT composites. Macromolecules **2002**, *35*, 9039.
37. Barraza, H.J.; Pompeo, F.; O'Rear, E.A.; Resasco, D.E. SWNT-filled thermoplastic and elastomeric composites prepared by miniemulsion polymerization. Nano Lett. **2002**, *2* (8), 797.
38. Park, C.; Ounaies, Z.; Watson, K.A.; Crooks, R.E.; Smith, J., Jr.; Lowether, S.E.; Connell, J.W.; Siochi, E.J.; Harrison, J.S.; St. Clair, T.L. Dispersion of single wall carbon nanotubes by in situ polymerization under sonication. Chem. Phys. Lett. **2002**, *364*, 303.
39. Wise, K.E.; Park, C.; Siochi, E.J.; Harrison, J.S. Stable dispersion of single wall carbon nanotubes in polyimide: The role of noncovalent interactions. Chem. Phys. Lett., submitted for publication.
40. Rao, A.M.; Eklund, P.C.; Bandow, S.; Thess, A.; Smalley, R.E. Evidence for charge transfer in doped carbon nanotube bundles from Raman scattering. Nature **1997**, *388*, 257.
41. Chen, R.J.; Zhang, Y.; Wang, D.; Dai, H. Noncovalent sidewall functionalization of single-wall carbon nanotubes for protein immobilization. J. Am. Chem. Soc. **2001**, *123*, 3838.
42. Débarre, A.; Jaffiol, R.; Richard. A.; Tchénio, P. Raman hyperspectral imaging applied to chemical co-localization in diluted samples of perylene-doped nanotubes. Chem. Phys. Lett. **2002**, *366*, 274.

43. McCarthy, B.; Coleman, J.N.; Czerw, R.; Dalton, A.B.; in het Panhauis, M.; Maitis, A.; Drury, A.; Bernier, P.; Nagy, J.B.; Lahr, B.; Byrne, H.J.; Carroll, D.L.; Blau, W.J. A microscopic study of interactions between carbon nanotubes and a conjugated polymer. J. Phys. Chem., B **2002**, *106*, 2210.
44. Coleman, J.N.; Curran, S.; Dalton, A.B.; Davey, A.P.; McCarthy, B.; Blau, W.; Barklie, R.C. Percolation-dominated conductivity in a conjugated-polymer–carbon–nanotube composite. Phys. Rev., B **1998**, *58* (12), R7492.
45. Dalton, A.B.; Stephan, C.; Coleman, J.N.; McCarthy, B.; Ajayan, P.M.; Lefrant, S.; Bernier, P.; Blau, W.J.; Byrne, H.J. Selective interaction of a semiconjugated organic polymer with single-wall nanotubes. J. Phys. Chem., B **2000**, *104*, 10012.
46. Chen, B.; Li, J.; Delzeit, L.; Partridge, H.; Meyyapan, M. Raman analysis of in-situ electropolymerized carbon nanotube and polypyrrole composite. Nano Lett., submitted for publication.
47. Mikat, J.; Orgzall, I.; Hochheimer, H.D. Raman spectroscopy of conducting polypyrrole under high pressure. Phys. Rev., B **2002**, *65*, 174202.
48. Furukawa, Y.; Tazawa, S.; Fujii, Y.; Harada, I. Raman spectra of polypyrrole and its 2,5-^{13}C-substituted and C-deuterated analogues in doped and undoped states. Synth. Met. **1988**, *24*, 329.
49. Jenden, C.M.; Davidson, R.G.; Turner, T.G. A Fourier transform-Raman spectroscopic study of electrically conducting polypyrrole films. Polymer **1993**, *34*, 1649.
50. Chen, J.H.; Huang, Z.P.; Wang, D.Z.; Yang, S.X.; Li, W.Z.; Wen, J.G.; Ren, Z.F. Electrochemical synthesis of polypyrrole films over each of well-aligned carbon nanotubes. Synth. Met. **2002**, *125*, 289.
51. Gao, M.; Huang, S.; Dai, L.; Wallace, G.; Gao, R.; Wang, Z. Aligned coaxial nanowires of carbon nanotubes sheathed with conducting polymers. Angew. Chem., Int. Ed. Engl. **2000**, *39*, 3664.
52. Chattopadhyay, D.; Galeska, I.; Papadimitrakopoulos, F. A route for bulk separation of semiconducting from metallic single-wall carbon nanotubes. J. Am. Chem. Soc. **2003**, *125*, 3370.
53. Krupke, R.; Hennrich, F.; v. Lohneysen, H.; Kappes, M. Separation of metallic from semiconducting single-walled carbon nanotubes. Science **2003**, *301*, 344.

Ring Structures from Nanoparticles and Other Nanoscale Building Blocks

Zhen Liu
Rastislav Levicky
Columbia University, New York, New York, U.S.A.

INTRODUCTION

Over the last two decades nanomaterials have attracted growing interest because of their unique, potentially useful electronic, magnetic, and optical properties. For example, carbon nanotubes possess tunable electronic structure from metals to insulators. Nanoparticles can act as "artificial atoms" which, when assembled into meso- or macroscopic structures may lead to novel functional materials. Thus self-assembled particle arrays not only preserve the properties of individual particles but can also exhibit new behaviors due to interparticle correlations and coupling. The ability to organize nanoscale components into specific geometrical arrangements, with defined interconnections, underpins much of their promise for future applications as well as current fundamental studies. Here we focus on the organization of nanoparticles and, to a lesser extent, other nanoscopic building blocks into ringlike arrangements ranging over five decades of length scale, from macroscopic ($\sim$1 mm) to nanoscopic ($\sim$10 nm). In addition to the simple geometry or a ring, a variety of other arrangements have been realized. For example, following earlier reports of close-packed crystalline nanoparticle superstructures,[1–5] Korgel et al. analyzed nanoparticle self-assembly into superlattices in detail highlighting the effects of size-selection, ligand coverage, and interparticle attraction.[6] Detailed reviews on the fabrication and properties of nanoparticle superlattices are now available.[7,8] Other efforts have focused on organizing nanoparticles according to highly specific biological interactions[9–19] or using self-organized polymeric media such as block copolymers to spatially template nanoparticle synthesis and distribution.[20–28] Albeit the focus of this review restricts it from broadly considering the general theme of nanomaterial self-assembly, the above and related efforts continue to make critical advances toward functional materials and devices incorporating nanoscale components.

The formation of ringlike assemblies, whether over macroscopic or nanoscale dimensions, is intriguing for reasons other than simple curiosity about their physical or chemical origin. Thus continuous rings, for example, will develop circulating currents when magnetic flux is applied across the ring.[29] This "persistent" current has a periodicity as a function of the enclosed magnetic flux. Moreover, the electronic states of the ring are likewise functions of the applied field, leading to magnetically tunable optical and other properties that, in addition to being fundamentally interesting, may provide unique capabilities in applications. While a variety of routes to microscopic rings have been reported, including lithography,[30] direct chemical synthesis,[31] and diffusive coalescence on free liquid surfaces,[32] self-assembly of ring structures using wetting and capillary phenomena has perhaps been the most common. By its nature, self-assembly often provides a facile fabrication route, and generally similar methods can be effectively used to organize different particle types, making these methods fairly general. In what follows, various mechanisms that have led to the formation of ring structures will be described.

SELF-ASSEMBLY OF MACROSCOPIC RINGS

Ringlike patterns formed by precipitated solute particles that are visible to the naked eye, such as coffee rings, are familiar phenomena. The formation of macroscopic ($\sim$1 mm to $\sim$1 cm) ring patterns from nonwetting droplets placed on a solid support was investigated by Deegan et al.[33] The mechanism presented by these authors invoked pinning of the contact line of the drying droplet as the carrier solvent evaporates. In order to maintain the contact line pinned at its initial position in the face of continued evaporation, a capillary flow of solvent is engendered from the interior of the droplet to its perimeter. Dispersed solute particles are carried by this resulting outflow to the edge where they accumulate (Fig. 1a). When the droplet fully dries, a ringlike residue consisting of the accumulated particles remains. Notably, in this mechanism the ring formation is largely independent of the nature of the underlying surface, the deposited solute particles, and the carrier solvent. Maenosono and collaborators formed millimeter scale annular rings by precipitating 4–6-nm CdS and CdSe nanoparticles from a suspension droplet (Fig. 1b).[34] These authors postulated that, in addition to the pinning of the droplet's contact line

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120013846

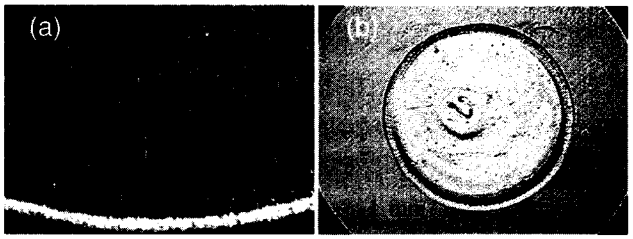


Fig. 1 (a) Motion of microspheres driven by the capillary flow of solvent from the interior to the periphery in an evaporating water droplet.[33] (b) Optical microscope photograph of a CdS ring with an outer diameter of 0.23 cm. (From Ref. [34].) (*View this art in color at www.dekker.com.*)

as discussed by Deegan et al., convection from the suspension droplet to its perimeter where ring assembly occurs could be driven by the capillary pressure from the meniscus curvature within the porous, drying ring of nanoparticles. In this model, the capillary pressure produces an effective pressure-driven suction of solvent into the ring. Both mechanisms are potentially operative in the formation of macroscopic rings of nanoparticles and provide useful insights into how such arrays may be assembled.

SELF-ASSEMBLY OF MICRON-SIZED RINGS

A number of investigators reported micron-scale annular ring assemblies of various nanometer subunits such as nanocrystals, nanotubes, and even small organic molecules. The origins of these microscopic structures are likely very different from those of macroscopic rings. Ohara and Gelbart observed micron-sized rings of nanoparticles,[35] possibly originating from the accumulation of particles along the rims of dewetting holes formed during the evaporation of nanoparticle suspensions on solid supports.[36] In these studies, the particle suspensions wetted the underlying support. Maillard and coworkers presented that thermocapillary flows engendered within an evaporating particle suspension could also lead to microrings and even hexagonal meshes of nanocrystals.[37,38] Stowell and Korgel observed the self-assembly of nanoparticles into rings and honeycomb networks.[39] The selection of a particular geometry was attributed to the variation of the thermal susceptibility of the suspension's surface tension because of the presence of nanoparticles, triggering a transition from hole nucleation in the evaporating film to Marangoni instability-induced convective flow. Microrings formed from other materials include a report by Schenning et al. of such structures generated from disk-like porphyrin molecules. These authors suspected that the rings formed as a result of hole nucleation possibly induced by foreign particles.[40] In the following discussion some of these mechanisms are examined in greater detail.

Thermocapillary Mechanisms

As discussed by Maillard et al. and others,[37–39] convective flow driven by variations in surface tension at the free interface of a thin evaporating liquid film can engender film rupture and hence the formation of holes or dry patches around which nanoparticles deposit. The physical origins of thermocapillary flows in liquid films have been well documented (e.g., see Ref. [41] for a review). For sufficiently thin (less than about 1 mm) volatile liquid films, evaporation generates a temperature gradient between the film's lower and upper boundaries due to the removal of latent heat, leaving the free upper surface cooler. Local fluctuations in the extent of cooling cause variations in the interfacial tension of this interface, with surface tension γ typically decreasing with temperature T according to a thermal coefficient $B = -d\gamma/dT$. As a result, the liquid from warmer surface regions where the interfacial tension is lower is pulled along the surface to spread over the cooler regions, where it is forced to move downward into the film as it cannot accumulate without limit. Simultaneously, warm liquid is fed upward to the warmer surface regions. Viscous coupling propagates this convective flow throughout the film to establish three-dimensional convective cells as illustrated in Fig. 2.

As evaporation proceeds, the thinning film may eventually rupture at the accompanying depressions, exposing the nearly bare substrate to form holes.[42–44] Evaporation and/or interfacial dewetting then drive the growth of the holes. When film rupture and subsequent hole growth occur in a suspension layer of nanoparticles, the particles can collect at the expanding hole rims eventually leaving behind drying patterns in the shape of rings (Fig. 3a). Possibly, the hole rims become pinned before complete drying of the film due to a build-up of frictional

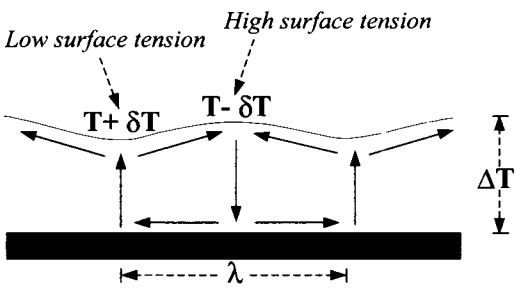


Fig. 2 Illustration of thermocapillary convective flows in volatile liquid films. The mean temperature drop across the film is ΔT. λ is a characteristic wavelength of the induced convective cells.

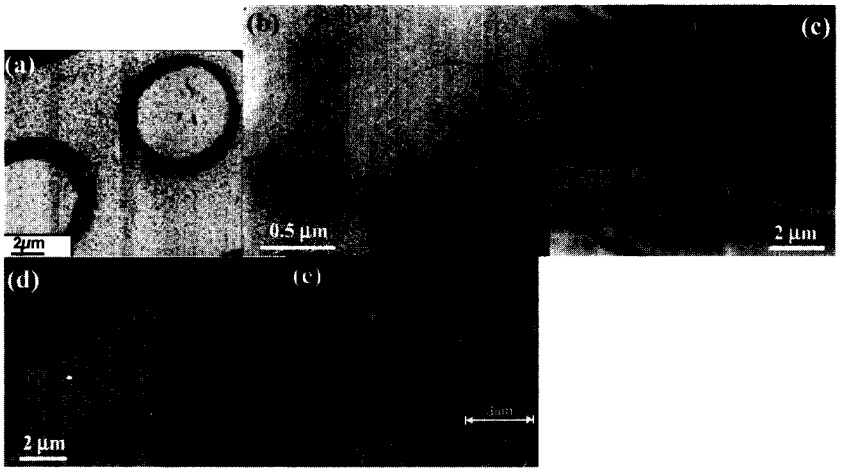


Fig. 3 (a) Microrings formed from ferrite nanocrystals precipitated out of hexane after deposition on a TEM grid under air.[37,38] (b) A microring of organically functionalized Ag nanocrystals precipitated out of hexane solution.[35] (c) Porphyrin rings generated from evaporating chloroform solutions.[40] (d) Microrings of gold nanocrystals prepared by spin coating from octane solutions on a polymer film (surface defects were indicated by red circles). (e) Nanotube rings formed after irradiation of bulk suspensions with ultrasound. (From Ref. [50].) (*View this art in color at www.dekker.com.*)

interaction between particles accumulated at the hole rim and the underlying surface.

A dimensionless Marangoni number, $M_a=(B\Delta Th)/(\rho\nu\kappa)$, is customarily used to compare the time scale for the propagation of temperature disturbances across the thickness of the film with that needed to establish convective flow.[45,46] Temperature perturbations leading to the establishment of convective flow must be sufficiently long lived, corresponding to a critical M_a value above which convective flow becomes possible, and below which Marangoni instabilities will not appear. The critical M_a depends on the boundary conditions at the film's interfaces.[45,46] At the most basic level, convective flow is favored by greater temperature variation across the film (ΔT), film thickness (h), and susceptibility of interfacial tension to variation in temperature (B), but is suppressed by increases in the liquid's kinematic viscosity (ν), density (ρ), or thermal diffusivity (κ).

Wetting Mechanisms

For apolar liquids and surfaces interacting purely via long-range dispersive forces, a drop of solvent placed on a solid support is typically characterized as either "wetting," meaning it thins so as to fully cover the support, or "nonwetting," meaning it beads up and forms a finite contact angle with the solid support. More completely, a set of classification rules can be devised that incorporate not only long-range apolar interactions but also short-range "contact" forces due to polar interactions.[47] By appropriate balancing of short polar and long-range apolar interactions, additional wetting states become possible including coexistence of thin wetting films with macroscopic, nonwetting droplets, and of dewetting thin films even when thicker layers of the same liquid wet. In considering the formation of microring assemblies from evaporating films, the approximation of either strictly (i.e., irrespective of film thickness) wetting or nonwetting liquids is usually sufficient as the surfaces and organic solvents used have been largely apolar in nature.

Ohara and Gelbart theoretically analyzed microring formation from evaporating particle suspensions when the suspension wets the surface.[35] Their analysis was motivated by accompanying experiments in which dilute nanoparticle solutions in wetting organic solvents such as hexane or toluene were spread on carbon-coated TEM grids and allowed to dry.[36] Drying patterns consisting of $\sim$1-μm-diameter rings were found (Fig. 3b). As a wetting liquid film evaporates, it will continue to thin until the film thickness decreases to nearly molecular dimensions, on the order of a few nanometers. At this stage, further thinning leads to increasing loss of attractive liquid–liquid dispersive interactions, which prefer to thicken the film. Ohara and Gelbart argue that, rather than dry uniformly to zero thickness, the film may seek to maintain a minimal thickness t_e to counteract the loss of favorable liquid–liquid interactions, following a previous theoretical analysis for nonvolatile liquid films by de Gennes.[48] In their model, as evaporation proceeds, the drying film will develop holes which will expand so as to maintain remnant wetted regions at the thickness t_e. The holes would open once the film thickness decreases sufficiently for free energy barrier to hole nucleation to become comparable to the thermal background kT, where k is the

Boltzmann constant. For typical solvents this occurs at thicknesses of ~ 1 nm. Once holes open and grow, nanoparticles can collect at the receding rims in a manner analogous to that described above for thermocapillary flows, leaving behind drying patterns in the shape of rings. Alternately, for sufficiently thin wetting films, holes can nucleate simply by spontaneous local evaporation of liquid molecules to form a dry patch or because of the presence of surface heterogeneities (e.g., a small non-wettable region).

Schenning et al. observed micron-sized ring assemblies composed of porphyrin molecules.[40] Their experimental procedure was similar to those used to form rings of nanoparticles. A droplet of porphyrin molecules in a solvent such as chloroform was placed on a substrate and allowed to evaporate for 10 sec, followed by draining of the remaining solution with filter paper. These authors suggested that the ringlike patterns originate from the nucleation of holes engendered by possibly combined influences of 1) nonwetability of the surface by the liquid; 2) progressive thinning of the liquid due to evaporation; and 3) deposition of foreign particles (e.g., from the laboratory ambient) that lowers the free energy barrier for hole nucleation. As the holes nucleate and grow, porphyrin molecules would accumulate at the rim increasing their concentration. Eventually, the concentration will exceed solubility and a condensed porphyrin phase will precipitate out to form the ring structures (Fig. 3c).

Liu and Levicky observed that spin coating of nanoparticles from organic solvents on polymer films produced micron-sized rings very similar to those reported by Ohara et al. and others.[35–40] Spin coating instead of droplet drying was used to better control deposition conditions. Samples for transmission electron microscopy (TEM) were prepared as reported previously.[28] From TEM micrographs, these authors identified that microrings formed around defects in the polymer film (Fig. 3d). Based on this evidence, ring formation was attributed to defect-induced nucleation of holes in the spin-coated particle solution followed by particle accumulation at the hole rims, leaving microring drying patterns as discussed above. The defects appear white in the images (Fig. 3d), suggesting that the polymer film is thinner (possibly punctured) at those locations. Others have similarly connected surface topography with nucleation of holes and subsequent formation of drying patterns. For example, step edge defects on highly oriented pyrolytic graphite surfaces led to the formation of ringlike patterns from drying collagen solutions.[49]

Microring assembly will, in general, also depend on other factors such as solute concentration and evaporation speed. Clearly, rings cannot form if the solute (e.g., nanoparticle) concentration is too low because of the insufficiency of available material. Moreover, microring formation, at least in wetting systems, appears favored by faster rates of evaporation. For instance, in experiments by the present authors in which nanoparticle suspensions were spin coated from relatively low volatility, wetting solvents such as octane, microrings did not form if the deposition speed was below 100 rpm or if static evaporation was used. Presumably, under such wetting conditions, a nucleated hole can quickly seal unless solvent evaporation from the hole rim is sufficient to ensure hole growth.

Other Types of Microrings

Microrings of carbon nanotubes have also been reported.[50,51] Martel and coworkers observed microrings of single-walled carbon nanotubes (SWNTs) following ultrasound irradiation of nanotubes in a warm solution of sulfuric acid and hydrogen peroxide (Fig. 3e). This process shortens the SWNTs and evidently can cause them to coil into rings. The physical origin of coiling was suggested to stem from tube bending due to the nucleation of bubbles on the hydrophobic surfaces of the SWNTs during ultrasound irradiation, followed by subsequent bubble collapse. Once a tube becomes coiled into a ring, the strain energy of bending can be balanced by physical van der Waals attraction along the nanotubes. This renders the rings highly permanent, sufficiently so that they can be recovered by filtering. High yields of rings, up to 50%, were reported.

SELF-ASSEMBLY OF NANOSCALE RINGS

Rings with diameters ranging from 10 to 100 nm have been made out of magnetic nanoparticles or metal nanocrystals by several groups. In contrast to micron-sized rings, these structures are approximately one particle wide along their perimeter (Fig. 4). The self-assembly mechanisms, as may be expected, are distinct from those effective for forming macro- or microrings and involve magnetostatic particle interactions in the case of magnetic

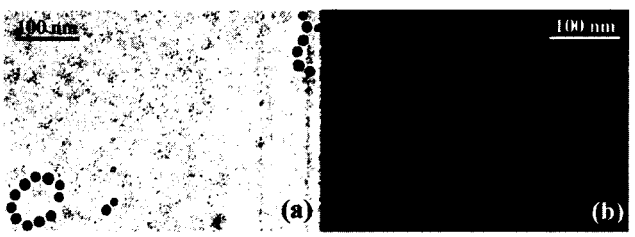


Fig. 4 (a) Nanorings formed by 16-nm cobalt nanoparticles.[13] (b) Nanorings composed of 2.5-nm gold nanocrystals, formed by spin coating from octane solutions.

nanoparticles.[52,53] The present authors have also assembled nanoscopic rings from metal nanoparticles by spin coating nanoparticle solutions on thin polymer films.

Magnetostatic Mechanisms

Tripp et al. synthesized weakly ferromagnetic cobalt nanoparticles (~30 nm in diameter), possessing a 3- to 0.4-nm oxide layer.[52] The nanoparticles were dispersed in toluene or CH_2Cl_2 in the presence of a resorcinarene as a stabilizing surfactant. Droplets of nanoparticle solution containing surfactant were deposited on carbon-coated TEM grids and allowed to dry. If the concentration of surfactant was sufficiently high, this procedure produced a large population of nanorings with diameter in the range of 50 to 100 nm. A possible mechanism for ring formation is minimization of magnetostatic energy realized when a string of magnetic dipoles (i.e., a string of "head-to-tail" aligned ferromagnetic Co nanoparticles) closes, thus eliminating unpaired magnetic poles at its ends. Notably, without sufficient surfactant the Co nanoparticles aggregated into dense, continuous layers or "rafts." The need for the nonvolatile resorcinarene surfactant in the assembly of rings was believed to stem from kinetic stabilization attributed to the enhancement of the viscosity of the deposited particle layer. Otherwise, at low surfactant concentrations, the nanoparticles were sufficiently mobile to rearrange into more thermodynamically stable, dense aggregates. Other aggregate structures, such as oriented strings of particles, were also identified with preference for a particular geometry somewhat adjustable by application of external magnetic fields.

A previous report by Puntes et al. similarly described, among many other structures, the self-assembly of nanorings from ferromagnetic Co nanoparticles (Fig. 4a).[53] Closed motifs, such as nanorings, did not form with smaller Co nanoparticles that were superparamagnetic rather than ferromagnetic. These results indicate that the permanency of the particles' magnetic dipole plays a crucial role in facilitating the formation of closed loops or rings. In the work of Puntes et al., the Co nanoparticles had a thin organic coating and were deposited onto carbon-coated TEM grids from evaporating suspensions.

Assembly of Nanorings on Polymer Films

Polymer media have been used to template spatial distributions of nanometer-sized particles in a variety of ways, including in situ synthesis of nanoparticles in ordered block copolymer matrices,[54–56] introduction of premade nanoparticles into the bulk or on the surface of such media,[25,57–61] or via additional routes exploiting polymeric micelles[62,63] or multilayered structures.[64–69] Liu and coworkers have been investigating the organization of nanoparticles in polymer "brushes,"[28] which are monolayers of polymer chains tethered by one end to a planar support. The polymer brushes consist of poly(ethylenepropylene) (PEP) chains and are typically between 5 and 20 nm in thickness. The PEP chains incorporated a shorter polystyrene (PS) endblock to provide anchoring to an underlying film of PS homopolymer. Dodecanethiol-stabilized gold nanoparticles about 2.5 nm in diameter were spin coated on top of the polymer brush from octane and their drying patterns imaged by TEM. On some specimens, rings of nanoparticles 10 to 50 nm in diameter were observed in high yields (Fig. 4b). The underlying physical cause of ring assembly has not yet been identified.

CONCLUSION

Fabrication of structures at submicron length scales is a difficult yet crucial capability in advancing devices and technologies that may one day incorporate nanomaterials. Prospective examples of applications include novel analytical tools, ultrasmall chemical sensors, optoelectronics, and new reagents and catalysts for use in chemistry and biology. The last few years have seen rapid developments in the self-assembly of a variety of structures from nanoparticles and other components such as nanotubes or nanorods. This brief review has specifically emphasized the methods for organizing tiny building blocks into ring-shaped assemblies, with the predominance of the responsible physical phenomena traced to wetting and hydrodynamics of thin liquid films. Even as the theme of self-assembly in nanomaterial science continues to expand, as illustrated by this report, at least for specific geometries general methods can begin to be formulated.

ACKNOWLEDGMENTS

During the writing of this work, Zhen Liu was supported by the MRSEC program of the National Science Foundation (DMR-0213574) and by the Donors of The Petroleum Research Fund, administered by the American Chemical Society.

REFERENCES

1. Murray, C.B.; Kagan, C.R.; Bawendi, M.G. Self-organization of CdSe nanocrystallites into three-dimensional quantum dot superlattices. Science **1995**, *270*, 1335.
2. Kagan, C.R.; Murray, C.B.; Nirmal, M.; Bawendi,

M.G. Electronic energy transfer in CdSe quantum dot solids. Phys. Rev. Lett. **1996**, *76*, 1517.
3. Harfenist, S.A.; Wang, Z.L.; Alvarez, M.M.; Vezmar, I.; Whetten, R.L. Highly oriented molecular Ag nanocrystal arrays. J. Phys. Chem. **1996**, *100*, 13904.
4. Taleb, A.; Petit, C.; Pileni, M.P. Synthesis of highly monodisperse silver nanoparticles from AOT reverse micelles: A way to 2-D and 3-D self-organization. Chem. Mater. **1997**, *9*, 950.
5. Andres, R.P.; Bielefeld, J.D.; Henderson, J.I.; Janes, D.B.; Kolagunta, V.R.; Kubiak, C.P.; Mahoney, W.J.; Osifchin, R.G. Self-assembly of a two-dimensional superlattice of molecularly linked metal clusters. Science **1996**, *273*, 1690.
6. Korgel, B.A.; Fullman, S.; Connolly, S.; Fitzmaurice, D.Assembly and self-organization of silver nanocrystal superlattices: Ordered "soft spheres." J. Phys. Chem., B **1998**, *102*, 8379.
7. Murray, C.B.; Kagan, C.R.; Bawendi, M.G. Synthesis and characterization of monodisperse nanocrystals and close-packed nanocrystal assemblies. Annu. Rev. Mater. Sci. **2000**, *30*, 545.
8. Collier, C.P.; Vossmeyer, T.; Heath, J.R. Nanocrystal superlattices. Annu. Rev. Phys. Chem. **1998**, *49*, 371.
9. Storhoff, J.J.; Mirkin, C.A. Programmed materials synthesis with DNA. Chem. Rev. **1999**, *99*, 1849.
10. Niemeyer, C.M. Self-assembled nanostructures based on DNA: Towards the development of nanobiotechnology. Curr. Opin. Chem. Biol. **2000**, *4*, 609.
11. Braun, E.; Eichen, Y.; Sivan, U.; Ben-Yoseph, G. DNA-templated assembly and electrode attachment of a conducting silver wire. Nature **1998**, *391*, 775.
12. Coffer, J.L.; Bigham, S.R.; Li, X.; Pinizzotto, R.F.; Rho, Y.G.; Pirtle, R.M.; Pirtle, I.L. Dictation of the shape of mesoscale semiconductor nanoparticle assemblies by plasmid DNA. Appl. Phys. Lett. **1996**, *69*, 3851.
13. Alivisatos, A.P.; Johnsson, K.P.; Peng, X.; Wilson, T.E.; Loweth, C.J.; Bruchez, M.P.; Schultz, P.G. Organization of "nanocrystal molecules" using DNA. Nature **1996**, *382*, 609.
14. Mirkin, C.A.; Letsinger, R.L.; Mucic, R.C.; Storhoff, J.J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. Nature **1996**, *382*, 607.
15. Taton, T.A.; Mucic, R.C.; Mirkin, C.A.; Letsinger, R.L. The DNA-mediated formation of supramolecular mono- and multilayered nanoparticle structures. J. Am. Chem. Soc. **2000**, *122*, 6305.
16. Torimoto, T.; Yamashita, M.; Kuwabata, S.; Sakata, T.; Mori, H.; Yoneyama, H. Fabrication of CdS nanoparticle chains along DNA double strands. J. Phys. Chem., B **1999**, *103*, 8799.
17. Niemeyer, C.M.; Burger, W.; Peplies, J. Covalent DNA–streptavidin conjugates as building blocks for novel biometallic nanostructures. Angew. Chem., Int. Ed. **1998**, *37*, 2265.
18. Willner, I.; Patolsky, F.; Wasserman, J. Photoelectrochemistry with controlled DNA-cross-linked CdS nanoparticle arrays. Angew. Chem., Int. Ed. **2001**, *40*, 1861.
19. Cassell, A.M.; Scrivens, W.A.; Tour, J.M. Assembly of DNA/fullerene hybrid materials. Angew. Chem., Int. Ed. **1998**, *37*, 1528.
20. Chan, Y.N.C.; Schrock, R.R.; Cohen, R.E. Synthesis of silver and gold nanoclusters within microphase-separated diblock copolymers. Chem. Mater. **1992**, *4*, 24.
21. Ciebien, J.F.; Clay, R.T.; Sohn, B.H.; Cohen, R.E. Brief review of metal nanoclusters in block copolymer films. New J. Chem. **1998**. *22*, 685.
22. Fogg, D.E.; Radzilowski, L.H.; Dabbousi, B.O.; Schrock, R.R.; Thomas, E.L.; Bawendi, M.G. Fabrication of quantum dot-polymer composites: Semiconductor nanoclusters in dual-function polymer matrices with electron-transporting and cluster-passivating properties. Macromolecules **1997**, *30*, 8433.
23. Zehner, R.W.; Lopes, W.A.; Morkved, T.L.; Jaeger, H.; Sita, L.R. Selective decoration of a phase-separated diblock copolymer with thiol-passivated gold nanocrystals. Langmuir **1998**, *14*, 241.
24. Spatz, J.P.; Mossmer, S.; Hartmann, C.; Moller, M.; Herzog, T.; Krieger, M.; Boyen, H.G.; Ziemann, P.; Kabius, B. Ordered deposition of inorganic clusters from micellar block copolymer films. Langmuir **2000**, *16*, 407.
25. Lopes, W.A.; Jaeger, H.M. Hierarchical self-assembly of metal nanostructures on diblock copolymer scaffolds. Nature **2001**, *414*, 735.
26. Thompson, R.B.; Ginzburg, V.V.; Matsen, M.W.; Balazs, A.C. Predicting the mesophases of copolymer–nanoparticle composites. Science **2001**, *292*, 2469.
27. Lee, J.Y.; Thompson, R.B.; Jasnow, D.; Balazs, A.C. Entropically driven formation of hierarchically ordered nanocomposites. Phys. Rev. Lett. **2002**, *89*. art. no.-155503.
28. Liu, Z.; Pappacena, K.; Cerise, J.; Kim, J.; Durning, C.J.; O'Shaughnessy, B.; Levicky. R. Organization of nanoparticles on soft polymer surfaces. Nano Lett. **2002**, *2*, 219.
29. Lorke, A.; Luyken, R.J.; Govorov, A.O.; Kotthaus, J.P.; Garcia, J.M.; Petroff, P.M. Spectroscopy of

nanoscopic semiconductor rings. Phys. Rev. Lett. **2000**, *84*, 2223.
30. Rabiei, P.; Steier, W.H.; Zhang, C.; Dalton, L.R. Polymer micro-ring filters and modulators. J. Lightwave Technol. **2002**, *20*, 1968.
31. Li, Z.J.; Chen, X.L.; Li, H.J.; Tu, Q.Y.; Yang, Z.; Xu, Y.P.; Hu, B.Q. Synthesis and Raman scattering of GaN nanorings, nanoribbons and nanowires. Appl. Phys., A **2001**, *72*, 629.
32. Yang, B.; Scheidtmann, J.; Mayer, J.; Wuttig, M.; Michely, T. Fragmentation, rings and coarsening: Structure and transformations of nanocrystal aggregate networks on a liquid surface. Surf. Sci. **2002**, *497*, 100.
33. Deegan, R.D.; Bakajin, O.; Dupont, T.F.; Huber, G.; Nagel, S.R.; Witten, T.A. Capillary flow as the cause of ring stains from dried liquid drops. Nature **1997**, *389*, 827.
34. Maenosono, S.; Dushkin, C.D.; Saita, S.; Yamaguchi, Y. Growth of a semiconductor nanoparticle ring during the drying of a suspension droplet. Langmuir **1999**, *15*, 957.
35. Ohara, P.C.; Gelbart, W.M. Interplay between hole instability and nanoparticle array formation in ultrathin liquid films. Langmuir **1998**, *14*, 3418.
36. Ohara, P.C.; Heath, J.R.; Gelbart, W.M. Self-assembly of submicrometer rings of particles from solutions of nanoparticles. Angew. Chem., Int. Ed. Engl. **1997**, *36*, 1078.
37. Maillard, M.; Motte, L.; Ngo, A.T.; Pileni, M.P. Rings and hexagons made of nanocrystals: A Marangoni effect. J. Phys. Chem., B **2000**, *104*, 11871.
38. Maillard, M.; Motte, L.; Pileni, M.P. Rings and hexagons made of nanocrystals. Adv. Mater. **2001**, *13*, 200.
39. Stowell, C.; Korgel, B.A. Self-assembled honeycomb networks of gold nanocrystals. Nano Lett. **2001**, *1*, 595.
40. Schenning, A.P.H.J.; Benneker, F.B.G.; Geurts, H.P.M.; Liu, X.Y.; Nolte, R.J.M. Porphyrin wheels. J. Am. Chem. Soc. **1996**, *118*, 8549.
41. Oron, A.; Davis, S.H.; Bankoff, S.G. Long-scale evolution of thin liquid films. Rev. Mod. Phys. **1997**, *69*, 931.
42. Burelbach, J.P.; Bankoff, S.G.; Davis, S.H. Nonlinear stability of evaporating condensing liquid-films. J. Fluid Mech. **1988**, *195*, 463.
43. Tan, M.J.; Bankoff, S.G.; Davis, S.H. Steady thermocapillary flows of thin liquid layers: I. Theory. Phys. Fluids, A **1990**, *2*, 313.
44. Burelbach, J.P.; Bankoff, S.G.; Davis, S.H. Steady thermocapillary flows of thin liquid layers: II. Experiment. Phys. Fluids, A **1990**, *2*, 322.
45. Pearson, J.R.A. On convection cells induced by surface tension. J. Fluid Mech. **1958**, *4*, 489.
46. Nguyen, V.X.; Stebe, K.J. Patterning of small particles by a surfactant-enhanced Marangoni–Bénard instability. Phys. Rev. Lett. **2002**, *88*, 164501.
47. Sharma, A. Relationship of thin film stability and morphology to macroscopic parameters of wetting in the apolar and polar systems. Langmuir **1993**, *9*, 861.
48. de Gennes, P.-G. Wetting: Statics and dynamics. Rev. Mod. Phys. **1985**, *57*, 827.
49. Thiele, U.; Mertig, M.; Pompe, W. Dewetting of an evaporating thin liquid film: Heterogeneous nucleation and surface instability. Phys. Rev. Lett. **1998**, *80*, 2869.
50. Martel, R.; Shea, H.R.; Avouris, P. Rings of single-walled carbon nanotubes. Nature **1999**, *398*, 299.
51. Liu, J.; Dai, H.; Hafner, J.H.; Colbert, D.T.; Smalley, R.E.; Tans, S.J.; Dekker, C.Fullerene ''crop circles.'' Nature **1997**, *385*, 780.
52. Tripp, S.L.; Pusztay, S.V.; Ribbe, A.E.; Wei, A. Self-assembly of cobalt nanoparticle rings. J. Am. Chem. Soc. **2002**, *124*, 7914.
53. Puntes, V.F.; Krishnan, K.M.; Alivisstos, A.P. Colloidal nanocrystal shape and size control: The case of cobalt. Science **2001**, *291*, 2115.
54. Chan, Y.N.C.; Craig, G.S.W.; Schrock, R.R.; Cohen, R.E. Synthesis of palladium and platinum nanoclusters within microphase-separated diblock copolymers. Chem. Mater. **1992**, *4*, 885.
55. Ciebien, J.F.; Clay, R.T.; Sohn, B.H.; Cohen, R.E. Brief review of metal nanoclusters in block copolymer films. New J. Chem. **1998**, *22*, 685.
56. Bronstein, L.; Seregina, M.; Valetsky, P.; Breiner, U.; Abetz, V.; Stadler, R. Transition metal complex induced morphology change in an ABC-triblock copolymer. Polym. Bull. **1997**, *39*, 361.
57. Fogg, D.E.; Radzilowski, L.H.; Dabbousi, B.O.; Schrock, R.R.; Thomas, E.L.; Bawendi, M.G. Fabrication of quantum dot-polymer composites: Semiconductor nanoclusters in dual-function polymer matrices with electron-transporting and cluster-passivating properties. Macromolecules **1997**, *30*, 8433.
58. Lin, B.H.; Morkved, T.L.; Meron, M.; Huang, Z.Q.; Viccaro, P.J.; Jaeger, H.M.; Williams, S.M.; Schlossman, M.L. X-ray studies of polymer/gold nanocomposites. J. Appl. Phys. **1999**, *85*, 3180.
59. Zehner, R.W.; Lopes, W.A.; Morkved, T.L.; Jaeger, H.; Sita, L.R. Selective decoration of a phase-separated diblock copolymer with thiol-passivated gold nanocrystals. Langmuir **1998**, *14*, 241.
60. Hamdoun, B.; Ausserre, D.; Joly, S.; Gallot, Y.;

Cabuil, V.; Clinard, C. New nanocomposite materials. J. Phys. II **1996**, *6*, 493.
61. Lopes, W.A. Nonequilibrium self-assembly of metals on diblock copolymer templates. Phys. Rev., E **2002**, *65*, 1063.
62. Spatz, J.P.; Roescher, A.; Moller, M. Gold nanoparticles in micellar poly(styrene)-*b*-poly(ethylene oxide) films-size and interparticle distance control in monoparticulate films. Adv. Mater. **1996**, *8*, 337.
63. Bronstein, L.; Kramer, E.; Berton, B.; Burger, C.; Forster, S.; Antonietti, M. Successive use of amphiphilic block copolymers as nanoreactors and templates: Preparation of porous silica with metal nanoparticles. Chem. Mater. **1999**, *11*, 1402.
64. Kotov, N.A.; Dekany, I.; Fendler, J.H. Layer-by-layer self-assembly of polyelectrolyte–semiconductor nanoparticle composite films. J. Phys. Chem. **1995**, *99*, 13065.
65. Schmitt, J.; Decher, G.; Dressick, W.J.; Brandow, S.L.; Geer, R.E.; Shashidhar, R.; Calvert, J.M. Metal nanoparticle/polymer superlattice films: Fabrication and control of layer structure. Adv. Mater. **1997**, *9*, 61.
66. Caruso, F.; Caruso, R.A.; Mohwald, H. Nanoengineering of inorganic and hybrid hollow spheres by colloidal templating. Science **1998**. *282*, 1111.
67. Joly, S.; Kane, R.; Radzilowski, L.; Wang, T.; Wu, A.; Cohen, R.E.; Thomas, E.L.; Rubner, M.F. Multilayer nanoreactors for metallic and semiconducting particles. Langmuir **2000**, *16*, 1354.
68. Gao, M.Y.; Richter, B.; Kirstein, S. White-light electroluminescence from a self-assembled Q-CdSe/PPV multilayer structures. Adv. Mater. **1997**, *9*, 802.
69. He, J.A.; Valluzzi, R.; Yang, K.; Dolukhanyan, T.; Sung, C.M.; Kumar, J.; Tripathy, S.K.; Samuelson, L.; Balogh, L.; Tomalia, D.A. Electrostatic multilayer deposition of a gold-dendrimer nanocomposite. Chem. Mater. **1999**, *11*, 3268.

Risk Assessment and Benefits

Douglas Mulhall
Marsh Harbour, Bahamas

INTRODUCTION

Risk assessment is the foundation of our capacity to evaluate economic, health, and ecological impacts. Yet, a wave of scientific discovery is transforming our understanding of not only risk but also the environments that sustain us. This will force basic changes to how we evaluate risk and make decisions to restrict new technology.

Until we acknowledge such changes and incorporate them into our assessment of risk, it will not be possible for detractors or supporters of nanotechnologies to accurately evaluate their risks and benefits.

OVERVIEW

Some critics perceive the risk of nanomaterials as too great to go ahead with until we study them further.[1] The most often-cited example is the risk that out-of-control nanomachines, which use carbon-based life forms for an energy source, will turn everything alive into "gray goo" as they multiply. Technology luminaries such as Eric Drexler and Robert Freitas,[a] Ray Kurzweil and Bill Joy,[b] along with science fiction writers such as Michael Crichton[c] have explored such potential threats. Related debates have often spilled over into the legislative arena as governments have limited development of biotechnologies such as stem cell research,[6] while initiating investigations into ways of restricting nanotechnology.[7]

These debates are similar to those presented in the early 1950s when it was postulated that nuclear proliferation would get out of control and lead to contamination of the whole world. As with the nuclear argument, critics say that this new generation of technologies may be too volatile to control. Similar arguments were made about the machine gun at the turn of the century. As every new major technology comes along, fears are expressed that it may annihilate humanity.

Many scientists disagree with these arguments and say that the potential benefits outweigh the potential risks.[d] They point to near-term benefits such as nanoscale methods that can detect and target diseases far more precisely and less invasively than we do today. Vast improvements in energy efficiency created by a new generation of solar cells may also solve our energy supply problems, they argue. They say further that while the destructive power of technologies has been growing enormously, this has not hindered the expansion of civilization on Earth so far.

So we are faced with a familiar quandary: risk the benefits or eliminate the risks?

Principles for controlling powerful technologies, so that we can manage risks and get benefits, have been around for some time. For example, as early as 1950 the well-known science fiction writer Isaac Asimov put forward his "Laws of Robotics."[9,10] Organizations that specialize in nanotechnology, such as the Foresight Institute have developed principles for managing nanotechnology risks.[11] Still other organizations such as the Center for Responsible Nanotechnology have drafted ethical guidelines to help cope with disruptive economic and social impacts.[12]

These works are each helpful. Together they may constitute the beginnings of a regulatory framework for administering technology risks. Yet by themselves and together they are still incomplete.

This article briefly describes three considerations that could transform the debate: *Technologies that merge with ecology, enhanced intelligence, and punctuated equilibrium.* Finally, the article shows how these discoveries may completely change the regulatory paradigm. This entry is necessarily limited on space and I encourage the interested reader to look at my book-length work *Our Molecular Future*[13] and the reference list for more details and information.

[a]The term "gray goo" was put forward by Eric Drexler in 1986 in Ref. [2] and has been extensively quoted and misquoted. For a summary of the gray goo discussion, see Ref. [3].

[b]Sun Microsystems co-founder Bill Joy[4a] revitalized the concept of voluntary relinquishment of dangerous technologies that was put forward by other scientists in relation to earlier technologies such as nuclear weapons. Inventor Ray Kurzweil disagrees with this approach. A summary of the discussion is found in Ref. [4b].

[c]Eric Drexler's 1980's description of uncontrollable nanobots was given great popularity by the science fiction novel: Ref. [5].

[d]Updates on the nanoethics debates, are found in the December and February newsletter of Ref. [8].

Dekker Encyclopedia of Nanoscience and Nanotechnology
DOI: 10.1081/E-ENN 120023668

BACKGROUND

Some scientists have begun to describe a point known as the "Singularity," where the rate of technology convergence makes it impossible for human beings to accurately forecast the near future. If this point is approaching, then it seems that all the present discussions by environmentalists, scientists, and government regulators over how to regulate nanotechnologies may soon become moot, because the process will be out of our hands regardless of what we do, unless we impose a draconian ban on all new technology development. Given our history, such a ban seems improbable; therefore we must consider the implications of exponentially accelerating technologies.

The ideas that *Homo sapiens*—as we are presently constituted—won't be able to control our own destiny, and that some other form of development that we can't comprehend right now may take over from present paradigms, seem so fatalistic that many people don't dare to consider such thoughts. These confound the basic human tenet of belief in the future. Yet with technology rapidly progressing, and with the particularly accelerating advances in nanotechnologies—that are described throughout this encyclopedia—we must consider such a possibility as a starting point in the discussion over how to regulate the risks posed by new technologies. Nor do we need to be fatalistic about it, because there are avenues whereby we can participate in the accelerated evolution that has begun to occur.

Just in the same way that the atomic bomb transformed the concepts of security and war, so advanced technologies are already upending conventional notions of evolution. For example, it has now become clear that computers have started to solve problems in ways that their human designers do not comprehend. Chess Grand Master Gary Kasparov acknowledged this reality many years ago when, after having been defeated by a computer at his own game, he stated that he had lost to "an alien."[14] This was rapidly followed by the development of genetic computing where software using genetic algorithms was used to design circuits in ways that human designers couldn't fully comprehend.[15]

Why is this so relevant right now to the discussion over regulation of nanotechnology?

The extrapolation to be drawn from this is that because some computers already exceed human intelligence in limited areas, that nanotechnology-enabled artificial intelligence will increasingly supercede our own. Given that it would be extremely difficult to regulate something that is smarter than a human, again we are confronted with this possibility that the regulatory discussion is moot.

However, such a viewpoint—and the fears expressed about "runaway" technology—overlook one central development. Human intelligence and machine intelligence are beginning to merge. As they do, the possibilities for anticipating and regulating development of further technologies take on a new light.

The perceptual problem that we face today is that most discussions over regulation of nanotechnologies are based on the assumption that while technology continues to evolve rapidly, *Homo sapiens* will somehow continue to evolve as we have for millennia: slowly and biologically. If this is true, then we can stop the discussion over regulation, because the rate of technological evolution is already exceeding the rate of biological evolution, and will render *Homo sapiens* intelligence obsolete or at least inferior.

However, this article looks at another possibility: that evolution of human intelligence is about to accelerate past our biological limitations into another realm. If so, this will transform the regulatory landscape.

Furthermore, our growing understanding of the natural environment is showing us that we may have no choice but to proceed, because history demonstrates that sooner or later nature will create conditions that make our existence on Earth difficult or untenable unless we take measures to protect ourselves.

NEW FACTORS THAT MAY TRANSFORM RISK ASSESSMENT

Using Life Cycle Assessment methodologies,[16–18] combined with an examination of new technological developments and new discoveries about the natural environment, these emerging theories and technologies can be identified as having the potential to profoundly transform the present paradigm of technology regulation.

Technologies That Merge with the Natural Environment

Ray Kurzweil, who pioneered technologies such as the flatbed scanner, argues that technology is a continuation of evolution by other means.[19] This implies that our technologies are becoming an integral part of the ecology. What are the physical manifestations of this?

Smart Dust[20] comprises a massive array of micromachines made of nanoscale components that ride on air or water currents, undetectable to the human eye. Each expendable machine can have a camera, communications device, and varying sensors for chemicals, temperature, and sound. It has its own rechargeable energy source. It can serve as the eyes, ears, nose, and guidance mechanism for everyone from soldiers to hurricane watchers. It may soon cost a fraction of a penny to manufacture, and its prototype exists today. It forms part of a massive array that delivers information to one or hundreds of computers

in one or many locations. It may soon be in our environment in the trillions, delivering information about everything from troop to sewage movements.

This nanoscale level of incursion into—and integration with—the ecology suggests emergence of an *intelligent environment*. Just as the natural environment exercises its own type of intelligence by passing on information from generation to generation via DNA, so we are creating an intelligent human-built environment not just alongside that, but also as part of it. An intelligent environment has elements that are able to sense virtually every part of the ecology, from the epicenter of earthquakes to the heart of a hurricane and the heartbeat of every species, then interpret what this means and how to react. Right now, we are only at the very first primitive stages of this, but our sensing capacities are accelerating.

Such intelligent particles are also gaining the capacity to *self-assemble*. Several universities have pioneered self-assembling photovoltaic materials that generate and conduct an electric current (see entry on "Photovoltaics for the Next Generation: Organic-Based Solar Cells").[21] These materials can be painted onto surfaces, thus eliminating the need for solar panels. Such chemical self-assembly is only a primitive precursor to molecular assembly that is described in other entries to this encyclopedia.

When we combine self-assembly with intelligent sensing at the nanometer scale, then multiply it a trillion-fold, we see that our technology is becoming an integral part of the ecology instead of just impacting it, and that human technologies may soon be indistinguishable from the natural environment. This is a profound transition.

Furthermore, such pervasive intelligence is developing outside the human brain, but also in deep contact with it.

Enhanced Intelligence Changes the Groundrules

Hans Moravec, of Carnegie Mellon's Robotics Institute, has shown convincingly—as have others—that the rate of acceleration in information processing is logarithmic.[22] Not only is the capacity to process ones and zeros multiplying, but the *rate* at which it is multiplying is also increasing.

For millennia, this exponential rate was barely perceptible, because it took thousands, then hundreds, then tens of years for such capacity to multiply, from the abacus to the microprocessor, and now the nanoprocessor.

Today, this exponential acceleration enables super-fast manufacturing by machines and software. An example of this is desktop manufacturing that is transforming desktop printing into three-dimensional desktop manufacturing of products.[23]

Such hyperchange is upending the ground rules for intelligence, and by extension for environmental risk management.

Most risk assessment today implicitly assumes that evolution of human intelligence will proceed in the same way that it has over the past few thousands of years—that is, gradually.

Here are examples of why this assumption may be wrong.

In 2001, a computer used "genetic computing"[24] to build a thermostat and actuator that were superior to the counterparts designed by a human. The computer's programmers were unable to trace how the computer reached its conclusion. This is because genetic algorithms allow computers to solve problems in their own way without human intervention.

Machines with enhanced intelligence do certain things far faster and better than we do. Not everything, but many things. Stockbrokers now use algorithms that forecast commodity markets more accurately than humans do.[25] Satellites that repair themselves and make unilateral data transmission decisions are already in orbit.[26]

Moreover, massive networks are enhancing our own intelligence. It is now possible for the layperson to perform Internet searches in real time to get answers to complex questions. This acceleration in data retrieval *by the general population* constitutes a mass enhancement to our own memories.

At the more specialized level, remote robotic surgery is creating a networked medical "mind" that can perform operations in and from many locations at once.[27]

The extraordinary development is that human intelligence and primitive forms of machine "intelligence" are merging already. This is apparent from the use of artificial retinas for the blind, where computer microchips are implanted in the eye then connected to the optical nerve. They interpret and relay visual information to the brain. The merging of human intelligence with genetic algorithms and massive networks is being applied to modeling of, for example, climate change, but it has only just begun to be applied to evaluation of phenomena such as those described below.

Understanding Punctuated Equilibrium

The theory of *punctuated equilibrium*[28] was first proposed in 1972 by Niles Eldredge and Harvard evolutionary biologist Stephen Jay Gould.[41] This holds that evolutionary change occurs relatively rapidly in comparatively brief periods of environmental stress, separated by longer periods of evolutionary stability. After many years of skepticism, their theory is now gaining acceptance. This is because proof is emerging.

Intelligent tools, such as those described earlier, are helping us to discover that the natural ecology experiences periodic instability that threatens our society; not just in the extended time frames that we used to think.

In 1994, Comet Shoemaker–Levy 9 (SL9) hit Jupiter,[29] blasting holes the size of Earth in its atmosphere. Had this hit the Earth, human life would have been virtually extinguished. Only recently have we developed the tools to see such distant impacts, and until such technologies were invented we could only theorize about how often catastrophic collisions occur.

Before that, it was thought that such upheavals happened only every few million years and that we'd have lots of time to see them coming. SL9 demolished this idea. It demonstrated that we live in a galaxy where life can be snuffed out on a planetary scale without warning—in this era, not just the distant past.

Furthermore, scientists have found that smaller events have upset the ecology here on Earth. Ice core and tree ring records show that around the year A.D. 536 an unknown event triggered a catastrophic cooling of the Northern Hemisphere, resulting in years without summers that led to wholesale crop failures and starvation.[30]

Thousands of samples taken from ice cores and tree rings around the world show that naturally induced climate flips occur more frequently than we once thought they do, and that they don't only unfold over centuries but also erupt in a few years.[31,32]

At the regional scale, in 1700, a fracture at the Cascadia subduction zone produced a gigantic tsunami that scoured much of the Pacific coast for miles inland, where many of our cities now stand.[33] In 1958, a 1,500-foot wave swept away a forest after a mountain collapsed into Lituya Bay, Alaska.[34]

At the nanometer scale we are also getting a surprise. Researchers have discovered vast numbers of nanoscale organisms that are a hundred times smaller than most bacteria. In geology, they are named nanobacteria,[35] nanobes, and nanoarchaea.[36] In human ecology, a similar-sized entity has been labeled *Nanobacterium sanguineum* or blood nanobacteria.[37] Despite the name, it may not be a bacterium at all, but instead seems to be a newly discovered infection with the unusual ability to form a tough shell consisting of the same type of calcium found in many diseases. For decades, researchers have seen evidence that epidemic illnesses such as heart disease are triggered by infection.[38] This was proven for stomach ulcers decades ago, but for other illnesses no one could find a culprit. Now it seems that one has been discovered,[39] as chronicled in the book *Has Heart Disease Been Cured?*[40]

If the existence of such organisms turns out to be valid after the hot discussion over them is resolved, then it may fundamentally alter our understanding of how ecology works, what constitutes an ecosystem, and how epidemics decimate populations.

The reality that has been overlooked by environmental agencies and theorists is that many of these nano- and macro-scale phenomena pose deep threats to our society. Agencies such as the United States Federal Emergency Management Agency (FEMA) and United States Environmental Protection Agency (EPA) have few defenses against them. Such agencies usually do not consider how to adapt to climate altering super-volcanoes or epidemics from strange nano-organisms, because these are perceived as indefensible, or they haven't entered the awareness of the organizations yet.

Thus, punctuated equilibrium is not part of the risk assessment framework. As such, a chunk of the equation is missing. This is especially true when considering the relative risks and benefits posed by nanotechnologies. Such technologies may be driving the next "punctuation" in evolution by upending longstanding paradigms. At the same time they may give us the tools to protect ourselves from newly discovered big natural threats. Such is the contradictory reality of the two-edged sword.

MATCH NATURE'S COMPLEXITY

The convergence of these discoveries may let us achieve something that we have only dreamed of until now: how to match nature's complexity.

Right now, most of our technologies are unable to match the complexity of natural environments. For example, we use antibiotics to cure bacterial infections, but they lose their potency when the environment that they work in adapts to them. We build power lines to survive ice storms, but our miscalculation of the worst scenarios leads to collapses that paralyze our high-technology infrastructures.

Most of our agricultural, medical, energy, transportation, and housing systems are in a constant struggle to respond to the complexity of the natural environment.

Yet this imbalance may shift. Molecular technologies are empowering us to find solutions that replicate natural processes at the molecular level (for some examples, see entries on "Biomedical Applications: Tissue Engineering, Therapeutic Devices, and Diagnostic Systems" and "Nanomaterials: New Trends."). We may see energy grids based on solar "paint" that slash the political and economic risks associated with fossil fuel infrastructures. Our drugs may be so precise that they backfire only occasionally instead of generating widespread immune responses as they do now.

This nascent capacity to match nature's complexity constitutes the next environmental revolution. Over the centuries, such efforts to replicate natural processes have

been criticized as arrogant and unachievable approaches to nature, and today are sparking a political and religious furor. Nonetheless, they may soon force us to redefine the boundaries of risk assessment.

WHAT TO DO

These new realities—enhanced intelligence, technologies that merge with the environment, and newly understood evolutionary paradigms—are the elephants in the room of risk assessment. To cope with them, we must initiate a new regulatory discussion. We must first acknowledge that the yardsticks for measuring risk are being moved dramatically by our own rapidly expanding knowledge.

Just a few small examples: Although nanoscale organisms have been identified in geological formations and the human body since the early 1990s, few projects have examined the implications for human or natural ecology or for environmental chemistry. NASA is studying it, as are the universities of Texas, McGill (Canada), Regensburg (Germany), Kuopio (Finland), Melbourne (Australia), and others. However, no major government initiative is considering the implications. At the opposite end of the scale, few if any governmental, environmental, or disaster preparedness agencies are examining newly discovered mega-scale anomalies such as the naturally induced climate flip of circa A.D. 536, or the giant west coast tsunami of 1700. These would certainly disrupt natural ecosystems and civilized society if they recurred today, and evidence suggests that they may.

No disaster preparedness or environmental agency yet examines how nanotechnologies might be used for adapting to such phenomena. Future initiatives may develop, but at this time they are not apparent. To rectify this, the author has suggested that a forum be held under the auspices of one or several of the nanotechnology Non-Governmental Organizations such as The Foresight Institute or Center for Responsible Nanotechnology, to examine such an issue.

Examples of technologies that might help us to adapt to "nature's time bombs," and to explosive risks that may be posed by nanotechnologies themselves, include:

- Artificially intelligent software that is transforming the way that we make products and carry out business. The role that artificial and enhanced intelligence will play in risk assessment is so far under-studied and overlooked. This takes us to the heart of the issue of evolution of human intelligence relative to thousands of years of biological evolution. Likewise, technologies that are merging with the human body and mind, such as artificial retinas and other implants, merit far greater attention, as these are the stepping stones to development of *Homo sapiens* who have enhanced evaluative capacities.
- Superstrong nanostructured materials such as aerogels that exist now and that may let human settlements withstand mega-hurricanes, earthquakes, and tornadoes without causing more environmental damage than they prevent. Furthermore, the self-assembling and disassembling properties of other newer materials may protect us from more serious near-earth object threats that are now considered impossible to defend against, and that have been badly misjudged as too infrequent to worry about.
- "Desktop manufacturing" that may replace thousands of polluting factories while producing materials such as self-assembling solar materials that may stabilize our energy supplies.
- Nanomedicine that is opening the doors to new solutions for many prevalent diseases, and that may stop epidemics that have retarded human progress for millenia.
- Many other technologies that are cited throughout this encyclopedia and constitute an excellent basis to start with.

Of equal importance is the process used to evaluate them. By expanding the interdisciplinary approach to technology, it is possible and necessary to bring together experts in fields that are infrequently combined. These include:

- Computer scientists who have applied artificially intelligent software to technologies that might be used for environmental adaptation and risk assessment.
- Biochemists, geologists, and physicians who discovered nanobacteria in the environment and human body, and also developed treatments that seem to reverse nanobacterial infections.
- Climatologists, geologists, and astronomers who uncovered evidence of recurring climate disruptions, giant tsunamis, and near-earth object collisions.
- Scientists who have developed adaptive technologies such as desktop manufacturing and self-assembling photovoltaic materials that may let us adapt rapidly to big ecological changes.
- Critics who have proposed moratoria on nanomanufacturing.

CONCLUSION

By focusing such wide-ranging expertise on the challenges depicted in this article, risk assessment could be made into a more effective tool for proponents, detractors, and users of advanced technologies.

CREDIT LINE

This essay is an adaptation of a paper "Reassessing Risk Management" by Douglas Mulhall, presented in the 2003 Proceedings of the World Future Society, WorldFuture 2003: 21st Century Opportunitites and Challenges: An Age of Destruction or An Age of Transformation, edited by Howard F. Didsbury Jr. (World Future Society, 2003). http://www.wfs.org/2003confvol.htm.

REFERENCES

1. *The Big Down, Atomtech Technologies Converging at the Nano Scale*; ETC Group, January, 2003. [online] www.etcgroup.org/documents/TheBigDown.pdf [May 15, 2003].
2. Drexler, E. *Engines of Creation: The Coming Era of Nanotechnology*; Doubleday: New York, 1986. [online] http://www.foresight.org/EOC/EOC_Chapter_11.html [October 22, 2003].
3. *Ecophagy and Gray Goo* 2002; Compiled by Robert A. Freitas Jr. [online] http://www.foresight.org/Nanomedicine/Ecophagy.html [October 23, 2003].

4a. Joy, B. Why the future doesn't need us. Wired Mag. **April 2000** (8.04). [online] <http:www.wired.com/wired/archive/8.04/joy.html>.

4b. Kurzweil, R. Promise and peril of the 21st century. CIO Magazine **Fall/Winter 2003** issue. [online] http://www.cio.com/archive/092203/kurzweil.html [October 23, 2003].

5. Crichton, M. *Prey*; Harper Collins: New York, 2003.
6. *European Parliament Bans Stem Cell Research*; Reuters, April 10, 2003. [online] http://www.medscape.com/viewarticle/452207 [October 23, 2003].
7. Oger, G. Nanotech riddle: How green is Europe's 'goo'? Small Times **August 14, 2003** [online] http://www.smalltimes.com/document_display.cfm?document_id=6495 [October 22, 2003].
8. CMP-Cientifica. www.cmpcientifica.com/cientifica/frameworks/generic/public_users/tnt_weekly/subscribe.htm.
9. Asimov, I. *I, Robot*; Genome Press: New York, 1950.
10. Asimov, I. *Robots and Empire*; Doubleday: New York, 1985. Updated.
11. *Foresight Guidelines on Molecular Nanotechnology*; Foresight Institute June 2000. [online] www.Foresight.org/guidelines/current.html [October 23, 2003].
12. Phoenix, C.; Treder, M. *Three Systems of Action: A Proposed Application for Effective Administration of Molecular Nanotechnology*; Center for Responsible. Nanotechnology, Revised November 2003. [online] <http://crnano.org/systems.htm> [December 25, 2003].
13. Mulhall, D. *Our Molecular Future: How Nanotechnology, Robotics, Genetics, and Artificial Intelligence Will Transform Our World*; Prometheus Books: Amherst, NY, 2002.
14. Chang, G. *World Chess Champion Loses Game 2 Against Computer 'Deep Blue' Amidst Media Frenzy*; Discovery: Canada, May 6, 1997. [online] http://www.exn.ca/Stories/1997/05/05/01.asp [October 22, 2003].
15. Yasunaga, M.; Kim, J.H.; Yoshihara, I. The Application of Genetic Algorithms to the Design of Reconfigurable Reasoning VLSI Chips. International Symposium on Field Programmable Gate Arrays archive Proceedings of the 2000 ACM/SIGDA Eighth International Symposium on Field Programmable Gate Arrays, Monterey, CA, United States, 2000; 116–125.
16. Braungart, M.; Engelfried, J.; Mulhall, D. *Criteria for Sustainable Development of Products and Production*; Fresenius Environmental Bulletin, Birkhauser Verlag: Basel, Switzerland, 1993; Vol. 2, 70–77.
17. Mulhall, D. Tools for adapting to big ecosystem changes. Futures Res. Q. **Fall 2000** *16* (3), 49–61.
18. Mulhall, D. Redefining Earth: A conceptual framework for nanoecology. Nanotechnol. Mag. **December 2001**.
19. Kurzweil, R. Are We Spiritual Machines? *The Material World: "Is That All There Is?"*; Future Positive, June 18, 2002. [online] futurepositive.synearth.net/2002/06/20.
20. *Smart Dust*; DARPA research Web site describing project on micro air vehicles, [online] robotics.eecs.berkeley.edu/~pister/SmartDust/ [March 15, 2003].
21. Solar Cells Go Organic. In *The Economist, Technology Quarterly*; July 20, 2002. [online] www.economist.com/science/tq/displayStory.cfm?story_id=1176099 [March 14, 2003].
22. Moravec, H. Robot. In *Power and Presence*; Oxford 1998; 60. "The number of MIPS in $1,000 of computer from 1900 to the present," [online] www.frc.ri.cmu.edu/~hpm/book98/fig.ch3/p060.html [March 15, 2003].
23. Personal fabrication on demand. Wired Mag. **April 9, 2001**, *9* (4).
24. Koza, J.R., (Stanford University); Bennett F.H., III (Genetic Programming Inc.); Andre D. (University of California, Berkeley); Keane M.A. (Econometrics Inc.) Chapter V. Genetic Programming III: Darwinian Invention and Problem Solving. In *Automated Synthesis of*

Analog Electrical Circuits; Morgan Kaufmann, 1999.
25. Cliff, D. *Artificial Trading Agents for Online Auction Marketplaces*; HP Labs: Bristol. [online] www-uk.hpl.hp.com/people/dave_cliff/traders.htm [March 15, 2003].
26. Satellite trio to test artificial intelligence software. Aviat. Week **May 30, 2001**. AviationNow.com, [online] www.aviationnow.com/avnow/news/channel_space.jsp?view=story&id=news/ssat0530.xml [March 15, 2003].
27. *FDA Approves New Robotic Surgery Device*; July 11, 2000. Food & Drug Administration news release. [online] www.fda.gov/bbs/topics/NEWS/NEW 00732.html.
28. Gould, S.J. Darwinian Fundamentalism. In *New York Review of Books*; June 12, 1997. [online] www.nybooks.com/articles/1151 [March 17, 2003].
29. Bruton, D. *Frequently Asked Questions About the Impact of Comet Shoemaker-Levy 9 with Jupiter*; Institute for Scientific Computation: February 2, 1996. [online] www.isc.tamu.edu/~astro/sl9/cometfaq2.html#Q3.1 [August 29, 2001].
30. Stothers, R.B. Mystery cloud of 536 A.D. Sci. Front. **May-June 1984,** (33). Reprinted from R.B. Stothers, "Mystery Cloud of A.D. 536," Nature, 307 (1984), 344, [online] www.science-frontiers.com/sf033/sf033p19.htm [August 12, 2001].
31. Alley, R.B. *The Two Mile Time Machine: Ice Core, Abrupt Climate Change, and Our Future*; Princeton University Press: Princeton, NJ, 2000.
32. Baillie, M.G.L. *A Slice Through Time: Dendrochronology and Precision Dating*; Routledge: London, 1995.
33. Nelson, L, A.R.; Atwater, B.F.; Bobrowsky, P.T.; Bradley, L.-A.; Claque, J.J.; Carver, G.A.; Darienzo, M.E.; Grant, W.C.; Krueger, H.W.; Sparks, R.; Stafford, T.W.; Stuiver, M. Radiocarbon evidence for extensive plate boundary rupture about 300 years ago at the Cascadia subduction zone. Nature **November 23, 1995**, *378*.
34. *The 1958 Lituya Bay Tsunami*; University of Southern California Tsunami Research Group. [online] www.usc.edu/dept/tsunamis/alaska/1958/webpages/index.html [March 17, 2003].
35. Folk, R.L. Nanobacteria: Surely not figments, but what under heaven are they? Nat. Sci. **March 4, 1997**. [online] naturalscience.com/ns/articles/01-03/ns_folk.html [January 25, 2003].
36. Huber, H.; Hohn, M.J.; Rachel, R.; Fuchs, T.; Wimmer, V.C.; Stetter, K.O. A new phylum of archaea represented by a nanosized hyperthermophilic symbiont. Nature **May 2, 2002**, *417*, 63–67.
37. Kajander, E.O.; Ciftcioglu, N.; Miller-Hjelle, M.A.; Hjelle, J.T. Nanobacteria: Controversial pathogens in nephrolithiasis and polycystic kidney disease. Curr. Opin. Nephrol. Hypertens. **2001**, *10*, 445–452.
38. Mawhorter, S.D.; Lauer, M.A. Is atherosclerosis an infectious disease? Clevel. Clin. J. Med. **May 2001**, *68* (5).
39. Rasmussen, T.E.; Kirkland, B.L.; Charlesworth, J.; Rodgers, G.P.; Severson, S.R.; Rodgers, J.; Folk, R.L.; Miller, V.M. *Electron Microscope and Immunological Evidence of Nanobacterial-Like Structures in Calcified Carotid Arteries, Aortic Aneurysms and Cardiac Valves*; March 6, 2002. Mayo Clinic and Foundation: Rochester, Minnesota; University of Texas, Austin.
40. Mulhall, D.; Hansen, K. *Has Heart Disease Been Cured?* The Writers' Collective: Rhode Island, 2003. www.calcify.com.
41. Eldredge, N.; Gould, S.J. *Punctuated Equilibra: An Alternative to Phyletic Gradualism*; Schopf, T.J.M., Ed.; Paleobiology; Freeman and Company: San Fransisco, 1972.

R

Index

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.

This is an inclusive index for all five volumes. Volume breaks are as follows: Volume 1: pages 1–892; Volume 2: pages 893–1796; Volume 3: pages 1797–2676; Volume 4: pages 2677–3296; Volume 5: pages 3297–3980.