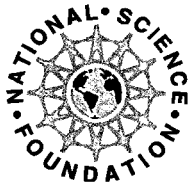
DOE/ER/14952



International Technology Research Institute
World Technology (WTEC) Division



NIST







WTEC Panel Report on

APPLICATIONS OF MOLECULAR AND MATERIALS MODELING

Phillip R. Westmoreland (Panel Chair)
Peter A. Kollman (Panel Vice Chair)
Anne M. Chaka
Peter T. Cummings
Keiji Morokuma
Matthew Neurock
Ellen B. Stechel
Priya Vashishta

January 2002

DOE Patent Clearance Granted
Mark P. Dvorscak
(630) 252-2393
E-mail: mark.dvorscak@ch.doe.gov
Office of Intellectual Property Law
DOE Chicago Operations Office
6-20-03
Date



International Technology Research Institute
R.D. Shelton, Director
Geoffrey M. Holdridge, WTEC Division Director and Series Editor

4501 North Charles Street
Baltimore, Maryland 21210-2699

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

WTEC PANEL ON APPLICATIONS OF MOLECULAR AND MATERIALS MODELING

Sponsored by the National Science Foundation, the Department of Energy, the National Institute of Standards and Technology, the Defense Advanced Research Projects Agency, the Air Force Office of Scientific Research, and the National Institutes of Health of the United States Government.

Phillip R. Westmoreland (Panel Chair)
Department of Chemical Engineering
University of Massachusetts Amherst
159 Goessmann, 686 N. Pleasant St.
Amherst, MA 01003-9303

Peter A. Kollman (Panel Vice Chair; deceased)
Department of Pharmaceutical Chemistry
University of California San Francisco
School of Pharmacy
San Francisco, CA 94143

Anne M. Chaka (previously at Lubrizol, Inc.)
National Institute of Standards and Technology
Physical and Chemical Properties Division (838)
Chemistry Building (222), Room A155
100 Bureau Drive, Stop 8380
Gaithersburg, MD 20899-8380

Peter T. Cummings
Department of Chemical Engineering,
University of Tennessee
Oak Ridge National Laboratory
University of Tennessee
Knoxville, TN 37996-2200

Keiji Morokuma
Department of Chemistry
Emory University
1515 Pierce Drive
Atlanta, Georgia 30322

Matthew Neurock
Department of Chemical Engineering
University of Virginia
Charlottesville, VA 22903-2442

Ellen B. Stechel
Ford Research Laboratory, Ford Motor Company
Scientific Research Laboratory, 2101 Village Road
Mail Drop 3083SRL/Rm. 3313
Dearborn, MI 48124-2053

Priya Vashishta
Departments of Physics & Astronomy
and of Computer Science
Nicholson Hall
Louisiana State University
Baton Rouge, LA 70803-4001

Special thanks to Sharon C. Glotzer (NIST, now at Univ. Michigan) Karl K. Irikura (NIST), Raul Miranda (NSF, now at DOE), Randall S. Jones (Loyola College), and Anna Tsao (Center for Computing Sciences), who contributed site reports to this volume.

INTERNATIONAL TECHNOLOGY RESEARCH INSTITUTE

World Technology (WTEC) Division

WTEC at Loyola College (previously known as the Japanese Technology Evaluation Center, JTEC) provides assessments of foreign research and development in selected technologies under a cooperative agreement with the National Science Foundation (NSF). Loyola's International Technology Research Institute (ITRI), R.D. Shelton, Director, is the umbrella organization for WTEC. Elbert Marsh, Deputy Assistant Director for Engineering at NSF's Engineering Directorate, is NSF Program Director for WTEC. Several other U.S. government agencies provide support for the program through NSF.

WTEC's mission is to inform U.S. scientists, engineers, and policymakers of global trends in science and technology in a manner that is timely, credible, relevant, efficient and useful. WTEC assessments cover basic research, advanced development, and applications. Panels of typically six technical experts conduct WTEC assessments. Panelists are leading authorities in their field, technically active, and knowledgeable about U.S. and foreign research programs. As part of the assessment process, panels visit and carry out extensive discussions with foreign scientists and engineers in their labs.

The ITRI staff at Loyola College help select topics, recruit expert panelists, arrange study visits to foreign laboratories, organize workshop presentations, and finally, edit and disseminate the final reports.

WTEC has now been spun off to a private, nonprofit corporation that will conduct all future WTEC studies, while continuing to assist in dissemination of older WTEC reports. See http://www.wtec.org.

Dr. R. D. Shelton
ITRI Director
Loyola College
Baltimore, MD 21210

Mr. Geoff Holdridge
WTEC Division Director
Loyola College
Baltimore, MD 21210

Dr. George Gamota
ITRI Associate Director
17 Solomon Pierce Road
Lexington, MA 02173

WTEC Panel on

APPLICATIONS OF MOLECULAR AND MATERIALS MODELING

FINAL REPORT

January 2002

Phillip R. Westmoreland (Panel Chair)
Peter A. Kollman (Panel Vice Chair)
Anne M. Chaka
Peter T. Cummings
Keiji Morokuma
Matthew Neurock
Ellen B. Stechel
Priya Vashishta

This document was sponsored by the National Science Foundation (NSF), the Department of Energy, the National Institute of Standards and Technology, the Defense Advanced Research Projects Agency, the Air Force Office of Scientific Research, and the National Institutes of Health of the U.S. Government under NSF Cooperative Agreement ENG-9707092, awarded to the International Technology Research Institute at Loyola College in Maryland. The government has certain rights to this material. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Government, the authors' parent institutions, or Loyola College.

ABSTRACT

This report reviews the development and applications of molecular and materials modeling in Europe and Japan in comparison to those in the United States. Topics covered include computational quantum chemistry, molecular simulations by molecular dynamics and Monte Carlo methods, mesoscale modeling of material domains, molecular-structure/macroscale property correlations like QSARs and QSPRs, and related information technologies like informatics and special-purpose molecular-modeling computers. The panel's findings include the following: The United States leads this field in many scientific areas. However, Canada has particular strengths in DFT methods and homogeneous catalysis; Europe in heterogeneous catalysis, mesoscale, and materials modeling; and Japan in materials modeling and special-purpose computing. Major government-industry initiatives are underway in Europe and Japan, notably in multi-scale materials modeling and in development of chemistry-capable *ab-initio* molecular dynamics codes. In European and U.S. assessments of nanotechnology, it was also concluded that to advance the field most quickly—and competitively—the need is acute for applying new and existing methods of molecularly based modeling. Additional findings are outlined in the panel's executive summary.

International Technology Research Institute (ITRI)

R. D. Shelton, Principal Investigator, ITRI Director

World Technology (WTEC) Division

(Staff working on this study)

Geoffrey M. Holdridge, WTEC Division Director and Series Editor
Bobby A. Williams, Financial Officer
Roan E. Horning, Head of Information Technologies

Aminah Grefer, Global Support Inc., Europe Advance Contractor
Gerald Whitman, ENSTEC, Inc., Japan Advance Contractor
Hiroshi Morishita, WTEC Japan Representative

IN MEMORY OF PETER A. KOLLMAN

1944-2001

The WTEC panelists, staff, and sponsors wish to express their profound sadness at the passing of panel Vice Chair Peter Kollman before the publication of this report. We wish to recognize Peter's lasting contributions not only to this report, but to the field of molecular modeling in general. His leadership of the pharmaceuticals applications side of this study, and of the study in general, was exemplary.

It was an honor to have Peter as a member of this panel and the WTEC extended family. We will miss him terribly.

Geoff Holdridge
WTEC Director

© 2002 to the printed work by Kluwer Academic Publishers except as noted herein. Copyright to electronic versions by Loyola College in Maryland and Kluwer Academic Publishers except as noted. This work relates to NSF Cooperative Agreement ENG-9707092. The U.S. Government retains a nonexclusive and nontransferable license to exercise all exclusive rights provided by copyright. This report is printed for the government under the above license. All WTEC reports are distributed on the Internet at http://www.wtec.org. A list of available JTEC/WTEC reports distributed on paper by the National Technical Information Service (NTIS) of the U.S. Department of Commerce, and information on ordering them, is included on the inside back cover of this report. This report will be published by Kluwer Academic Publishers in hardback form later in 2002.

TABLE OF CONTENTS

Table of Contents i
List of Figures v
List of Tables vi

Executive Summary vii

1. **Introduction**
Phillip R. Westmoreland

Goals and General Approach 1
Reasons for the International Study 2
Scientific and Technological Scope 3
Scope Beyond the Science of Molecular Modeling 4
Organization of the Analysis 5

2. **Science: Electronic Structure, Thermochemistry and Kinetics**
Keiji Morokuma

Introduction 7
Semi-empirical MO Methods 8
Ab Initio Methods 8
Electronic Density-Functional Theory (DFT) 11
Hybrid Methods 12
Thermochemical Calculations 13
Kinetics Calculations 14
Summary 15
References 15

3. **Science: Molecular Simulations and Mesoscale Methods**
Peter T. Cummings

Introduction 17
Molecular Simulation Methods 19
Issues 25
Acknowledgement 29
References 29

4. **Science: Information Technologies**
Peter A. Kollman

Introduction 35
Computer Graphics 35
Distance Geometry 35
QSAR/QSPR 35
Using Simulations in Structure Determination and Refinement 36
Docking into Biological Targets 36
Bioinformatics and Chemical Informatics 37
Computer Science 37
References 37

5. **Applications: The Chemical Industry**
Anne M. Chaka

Introduction 39
Chemical Industry Trends and Drivers: Impact of and on Modeling 40
What is the Role of Molecular and Materials Modeling in the Chemical Industry? 41
Needs of the Chemical Manufacturing Industry 42
Present State of the Art: Methods 43
Applications 45
How Modeling Technology is Transferred to Industry 46
Challenges and Needs 46
Summary 49
References 49

6. **Applications: Pharmaceuticals and Life Sciences**
Peter A. Kollman

Introduction and Historical Overview 51
The Approaches to Molecular Modeling in the Pharmaceutical and Biotechnology Industries 52
Structure and Duties of the Modeling Groups in Pharmaceutical and Biotech Companies 53
Hardware and Software in Modeling Groups in Pharmaceutical Companies 53
Relationship of Modeling to Other Technologies 54
Overview of Role, Successes and Failures of Modeling Groups 55
Needs for the Future 55
International Competitiveness in the Pharmaceutical Industry 56
References 57

7. **Applications: Physical and Electronic Materials**
Priya Vashishta

Introduction 59
Methods 60
Applications 63
Future Directions 67
Comparison of Science (Materials Simulations of Metals, Semiconductors, Ceramics, and Carbon Systems) and Industrial Application of Science in the United States, Europe and Japan 70
Conclusions 74
References 74

8. **Applications: Catalysis**
Matthew Neurock

Introduction 77
Methods 79
Present Applications 85
Panel Findings 94
Needs and Limitations 99
Future Directions 103
References 104

9. **Infrastructure Issues for Applying Molecularly Based Modeling**
Ellen B. Stechel

Introduction 107
Problem Solving 108
The Meaning of Successful Application 109
Necessary Attributes for Success 110
Closing Remarks 113

10. **International Comparison**
Phillip R. Westmoreland and Peter A. Kollman

Introduction 115
International Activities in Molecularly Based Modeling 115
Government Activities in Molecularly Based Modeling 147
Non-Government, Multi-lateral Activities 158
Summary of International Comparison 159

11. **Summary and Vision**
Phillip R. Westmoreland

Overall Conclusions 163
Major Findings: Applications 163
Major Findings: Process of Effective Application 166
Vision 167
References 171

APPENDICES

A. **Biographical Sketches of Panel Members and other Traveling-Team Members** 173

B.1. **Site Reports—Europe**

Air Liquide 180
BASF AG 183
Bayer AG 188
Belgium: University of Antwerp, Janssen Pharmaceutica, Solvay 191
BG Technology Ltd. (now Advantica Technologies Ltd.) 194
Center for Atomic-Scale Materials Physics (CAMP Denmark) 196
Centre Européen de Calcul Atomique et Moléculaire (CECAM) 199
Centre National de la Recherche Scientifique (CNRS) 201
Daresbury Laboratory 204
Degussa-Hüls AG (now Degussa) 213
DSM Research 215
EniChem S.p.A. 219
EniTecnologie SpA 222
GdR 12090 Project 224
Glaxo Wellcome plc (now GlaxoSmithKline) 226
Hoffmann-La Roche Ltd. 228
Novartis Pharmaceuticals AG 228
IBM Zürich Research Laboratory 231
Institut Français du Pétrole (IFP) 235
The Netherlands, including TU Eindhoven, Royal Dutch Shell, Philips 241
Rhône-Poulenc Industrialisation SA (now Rhodia) 247
Royal Society of Chemistry Molecular Modeling Group 250
SmithKline Beecham Pharmaceuticals (now GlaxoSmithKline) 256
TotalFina (now TotalFinaElf) 258
Unilever 260

B.2. **Site Reports—Japan**

Asahi Chemical Industry Co., Ltd. 261
"The Doi Project" Nagoya University 264
Eisai Co. Ltd. Tsukuba Research Laboratories 268
Fujitsu Ltd. and Fujitsu Laboratories 270
Institute for Molecular Science (IMS) 275
Institute for Solid State Physics (ISSP) 280

Japanese government organizations funding molecular and material modeling: 286
Joint Research Center for Atom Technology (JRCAT) 288
Angstrom Technology Partnership (ATP) 288
National Institute for Advanced Interdisciplinary Research (NAIR) 288
Tsukuba Advanced Computing Center (TACC) 288
Mitsubishi Chemicals, Yokohama Research Center 292
NEC Tsukuba Laboratories 296
Nagoya University 301
RIKEN—The Institute of Physical and Chemical Research 303
Taisho Pharmaceutical Co., Ltd. 307
Takeda Chemical Industries Ltd. 309
Toshiba Corporation, Research & Development Center 311
Toyota Central Research & Development Laboratories, Inc. 314
Short reports from other Japanese companies:
Fuji Photo Film Co., Ltd.
Japan Polyolefins, Ltd. (Showa Denko K.K.)
JSR Corporation
Mtsui Chemicals
Nippon Zeon Co., Ltd.
Sumitomo Chemical Co., Ltd.
UBE Industries, LTD. 315

B.3. **Site Reports—United States**

3M, Inc. 318
Air Products and Chemicals, Inc. 322
BP p.l.c. (Amoco) 327
CambridgeSoft.com 337
Chevron Corporation 339
The Dow Chemical Company 341
Dow-Corning Corporation 345
DuPont Pharmaceuticals (now part of Bristol-Myers Squibb) 348
E. I. du Pont de Nemours and Company 352
Eastman Chemical Company 356
Ford Motor Company 358
General Electric Corporation 363
HRL Laboratories, LLC 365
Louisiana State University 367
Lucent Technology / Bell Laboratories 374
Marathon Oil Company 378
Merck & Co., Inc. 379
Molecular Simulations, Inc. (MSI; now Accelrys) 382
Motorola, Inc. 387
National Institute of Standards and Technology (NIST) 390
Owens Corning, Inc. 393
Pharmacopeia, Inc. 394
Phillips Petroleum Company/Chevron Phillips Chemical Company 397
Rohm and Haas Company 400
Sandia National Laboratories 403
Solutia Inc. 408
University of Minnesota Supercomputing Institute 410

C. **Observations on Funding Estimates for Molecular Modeling** 415

D. **Glossary** 419

E. **Index of Companies** 426

LIST OF FIGURES

2.1. Stationary points on a potential-energy surface. 7
2.2. Hierarchy of methods and basis functions to approach a full, exact solution to the Schrödinger equation 9
2.3. Illustration of the concept of a conical intersection 10

3.1. Molecular modeling methods 18
3.2. Schematic of NEMD simulations of planar Couette flow 22
3.3. Schematic of GEMC 23

7.1. Results of a DD simulation of plastic flow localization in irradiated copper 64
7.2. Si(001)-SiO_2 interface structure from first-principles molecular dynamics. 65
7.3. An (8,0) C nanotube before collapse, and after collapse 67
7.5. "Moore's law" for molecular dynamics. The number of atoms in classical MD simulations (circles) has doubled every 18 months in the past 36 years, while that in quantum mechanical MD simulations (squares) has doubled every 13 months in the past 15 years. Petaflop computers anticipated to be built in the next ten years should maintain these rates. 68
7.6. A scientist immersed in an atomistic model of a fractured ceramic nanocomposite 69

8.1. Illustration of single-site metallocene complex 79
8.2. Structure of different forms of crystalline catalytic materials 80
8.3. Comparison of calculated and experimental chemisorption energies and vibrational frequencies for surface adsorbates 81
8.4. Three approaches and examples for modeling chemisorption and reactivity on surfaces 81
8.5. The elementary steps in the classic Cossee-Arlman mechanism for olefin polymerization over single-site catalysts 85
8.6. Activation barrier predictions for CHx insertion for olefin polymerization 86
8.7. Metal design for metallocene olefin polymerization catalysis. 87
8.8. Ligand structures used in the modeling approach of Union Carbide 88
8.9. Current simulation software is used to predict structure as well as sorption properties in zeolites 89
8.10. Overall catalytic cycle for methanol to methyl-tertiary-butyl-ether catalyzed by acid catalysis in a zeolite 91
8.11. Activation of N_2 on the Ru(0001) terraces and step edges 91
8.12. The effect of gas-phase oxygen partial pressure on the stability of the surface structure 92
8.13. Snapshot of the DFT-based kinetic MC simulation for the synthesis of vinyl acetate from ethylene and acetic acid over PdAu alloy 93

LIST OF TABLES

ES.1 Comparative Status of Molecular and Materials Modeling in Europe, Japan, and the United States ix

5.1 Examples of Companies Involved in the Chemical Industries 40
5.2 Comparison of Necessary Problem-Solving Skills in Industry and Academia. 42
5.3 Properties Accessible to Molecular Modeling and Simulation Techniques (properties that may be required in process design are highlighted) 44

8.1 Critical Areas for Molecularly Based Modeling of Catalysis 78
8.2 Summary of the Basic Molecular Simulation Methods Used in Modeling Homogeneous and Heterogeneous Catalysis 84
8.3a European-based Activities In Molecularly Based Catalysis Modeling, Industrial and Government Sites 94
8.3b Japanese-based Activities in Molecularly Based Catalysis Modeling, Industrial and Government Sites 95
8.3c U.S.-based Activities in Molecularly Based Catalysis Modeling, Industrial and Government Sites 95
8.4 General Industrial Computational Needs 100

10.1 Recent and Ongoing Modeling Activities in Europe 116
10.2 Recent and Ongoing Modeling Activities in Japan 133
10.3 Modeling Activities in Africa, Australia, and Asia Outside Japan 138
10.4 Recent and Ongoing Modeling Activities in North and South America Outside the United States 140
10.5 Recent and Ongoing Commercial Modeling Activities in the United States 142
10.6 Estimates of NSF Funding Involving Molecularly Based Modeling, FY98 ($ million) 157
10.7 Comparative status of Molecular and Materials Modeling in Europe, Japan, and the United States 160

EXECUTIVE SUMMARY

Phillip R. Westmoreland[1]

INTRODUCTION

From its purely theoretical origins, molecularly based modeling has moved to extensive practical applications. Molecularly based modeling here refers to the following:

- Computational quantum chemistry
- Molecular simulations by molecular-dynamics and Monte Carlo methods
- Mesoscale modeling of material domains
- Molecular-structure / macroscale property correlations like QSARs and QSPRs
- Related information technologies like informatics and special-purpose molecular-modeling computers

This report finds strong evidence that in the next ten years, these approaches will profoundly affect how chemistry, biology, and materials physics are understood, communicated, and transformed to technology, both in intellectual and commercial applications. In the past five to ten years, such computational approaches have taken hold. Growth in applications has been slow and sometimes fitful, often for nonscientific reasons. However, these approaches have brought powerful new means to understand chemical reactions and materials physics, and they are playing a crucial role in orienting applied life sciences toward biology's chemical underpinnings.

Indeed, one sign of impact is that specialists have now been joined by general users of these methods. Whether in a journal, a classroom, or a boardroom, presentations of three-dimensional molecules have become widespread and can be extremely effective. They are made possible by simple methods of generating and representing them. Similarly, new employees are used to having and using basic molecular-modeling software. As a linked consequence, experts can be more effectively employed in tasks that demand their particular skills, especially when general users understand both their own capabilities and limitations.

This study focuses on commercial applications of molecularly based modeling, examining successes, failures, and the causes of each. Advances depend not only on the right science, but also in the way the science is applied. For industry or anyone else to choose these methods, useful impact has to be proved. If new tools seem valuable, people rush to adopt them. If new tools fail, people rush away from them, often failing to ask why. It is interesting that most failures of molecular modeling tools prove to be based ultimately on non-technical causes. The process of application then is a vital part of successful application.

Even in an era of instant global communications, there are clear regional strengths. These strengths are rooted in geographical centers of intellectual achievement. The origin of these centers is individuals and small teams of gifted researchers. These strengths have evolved because government, industry, or individual commitment to specific ideas and approaches has fostered intellectual centers, and they expand regionally because technology is transferred best by the movement of knowledgeable people. Each region has benefited by people spending careers or extended periods across international borders, but Japanese organizations have

---

[1] Representing the consensus of all the panel members.

been among the most conscientious and successful about doing so. European countries have been especially effective at creating multi-national research and educational programs.

These observations and conclusions are based on the expertise of an eight-person panel and on information they collected in a 1999-2001 study. In February 1999, a team of Federal agencies proposed a study to evaluate how molecularly based modeling is being applied throughout the world, building on previous U.S.-focused studies and giving special attention to where U.S. activities stand. Requesting agencies included the National Science Foundation (Directorate for Engineering, Directorate for Mathematical and Physical Sciences/Division of Materials Research), Department of Energy (Basic Energy Sciences/Office of Science, Office of Industrial Technology), Department of Defense (DARPA and AFOSR), National Institutes of Health, and National Institute of Standards and Technology. As a key part of the study, the panel made 53 site visits in Europe, Japan, and the United States, mostly to commercial organizations, choosing to rely on personal knowledge and the open literature for including academic and government activities. Detailed reports were prepared on activities at a total of 91 organizations (see Appendix B). The study was managed by the World Technology (WTEC) Division of the International Technology Research Institute, based at Loyola College in Baltimore.

The following chapters and the case-study reports (Appendix B) give details of the science and technologies, as well as an evaluation of international activities in this field. Related funding of these activities by the United States government is estimated at $200 million, as summarized in Appendix C. Also, a glossary is provided in Appendix D.

Findings may be grouped into two categories: (1) applications and (2) the process of effective application. Table ES.1 presents a distillation of this panel's opinions, which are cheerfully recognized to be imperfect. While the members of this panel recognize that there are significant activities in other parts of the world, the table restricts itself to reflections upon activities in the U.S., Europe, and Japan. Certain broad findings stand out, and they are summarized here.

FINDINGS: APPLICATIONS

(1) Molecular modeling methods have gained acceptance as practical tools in a variety of industries.

From modest beginnings, typically a single specialist in a technical services or research role, many companies have developed effective molecular modeling activities. Noting a few example companies, activities include development of products and processes involving the following:

- Bio-active materials like pharmaceuticals (Merck, Novartis, Takeda Chemicals) and crop-protection chemicals (DuPont, Sumitomo Chemical)
- Polymers, glass, and structural materials (Asahi Chemicals, Owens-Corning, Rhône Poulenc / Rhodia, W.R. Grace)
- Electronic and photonic materials (Motorola, Toshiba, Lucent)
- Homogeneous and heterogeneous catalysts (Ford, Haldor Topsøe, Ube Industries)
- Sorbents for gas separations (BG Technologies / Advantica Technologies Ltd., Air Liquide, Air Products and Chemicals)
- Personal-care, food, and consumer products (Colgate Palmolive, Unilever, Kellogg, 3M)
- High-volume chemicals and materials (Dow, BASF, Rohm & Haas)
- Dyes and pigments (Bayer, Mitsubishi)
- Films and imaging (Fuji Photo Film, Xerox)
- Fuels and automotive chemicals (Chevron, TotalFina / TotalFinaElf, Lubrizol)
- Commercial software and hardware for calculations (MSI / Accelrys / Pharmacopeia, Gaussian, COSMOlogic, Fujitsu)

Table ES.1
Comparative Status of Molecular and Materials Modeling in Europe, Japan, and the United States

Technology	Europe Status	Japan Status	U.S. Status
Methods			
- Quantum Mech & Hybrid Methods, Molecular	****	***	*****
- Quantum Mech. & Hybrid Methods, Condensed Matter	*****	****	****
- Molecular Simulations			
Biological systems	****	***	*****
Fluids	*****	***	*****
Solids, devices	***	****	***** (devices)
Polymers	*****	*****	****
Kinetics – Monte Carlo	***	**	***
Mesoscale	*****	***	***
- Solvation	*****	****	*****
- Informatics & Optimization	****	***	*****
- Hardware, Operating Systems, Math Methods	****	*****	*****
Applications			
- Chemical Process Development	****	****	****
- Chemical Process Design and Manufacturing	**	**	**
- Chemical Product Design Fuels, specialty chemicals, coatings, surfactants	****	****	****
- Reaction Thermochemistry & Kinetics			
Noncatalytic	****	***	*****
Homogeneous catalysis	***	*****	*****
Heterogeneous catalysis (mostly academic)	**** (QC)	***	**** (sim)
- Physical and Electronic Materials			
Metals, semiconductors	***	*****	*****
Ceramics, glasses	****	*****	****
Polymers	*****	****⇑	*****
- Biomolecules and Biologically Active Materials			
Pharmaceuticals	****	***	*****
Agricultural products	***⇑	?	***⇑
- Application to Analytical Measurements			
Chemical analysis	***⇑	?	****⇑
Material analysis	***⇑	**⇑	**⇑
Biological analysis	****	***	*****
Transfer of Science to Technology	****	*****	***

Key: # of stars indicates level and quality of activity in each region;
0 = no activity, no quality; 5 = very large amount of activity or high level of quality;
up arrows (⇑) indicate increasing level of activity

(2) The methods are being adopted because they are proving their value in big ways.

Big successes naturally attract the most attention, both from within companies and from competitors. The study identified three major success areas:

- *Drug discovery* has been the most prominent success. Docking calculations and modeling of protein crystallography by molecular simulations have been important in developing both promising leads and

successful drugs. Commercial examples are the HIV protease inhibitor Agenerase™, discovered by Vertex Pharmaceuticals (Cambridge MA) using structure-based drug design (FDA approval 1999), and Aricept®, an ACE inhibitor used as an Alzheimer's disease medication from Eisai Co., Ltd. (a Japan-based global company).

- The profound impact on developing *homogeneous catalysts* has been less well known but has been nearly as significant. In 1975, polyolefins were projected to fall to 20% of the polymers market by 1995; instead, in 1995 they were at 85%, largely because of improvements in polymer properties due to homogeneous catalysts. Promising leads and commercial catalysts have been developed or screened by computational quantum chemistry at a large number of companies, including Asahi Chemical, BASF, Bayer, BP (Amoco), Degussa, Dow Chemical, DSM, DuPont, Enichem, IFP, Mitsubishi, Phillips Petroleum, Statoil, Totalfina, and Union Carbide.
- *Computational thermochemistry* has proved to be a key deliverable in many companies, mostly for ideal-gas thermochemistry. These numbers are widely used in evaluation of reaction equilibria, process design, and process safety analysis. In a 1996 conference presentation, a Dow Chemical speaker told how the heat of formation for a molecule of interest would cost about $70,000 to be measured calorimetrically but $20,000 for a value of comparable accuracy by a G2 quantum chemistry calculation. At the start of 2000, their comparison had changed to $100,000 or more for the experimental approach versus $2,000 for G3 calculation with comparable accuracy.

(3) Many companies now see such modeling as necessity, while others still consider it a luxury.

Management and employees may view these activities as necessary, on trial, or merely of possible long-range value—a luxury that can be eliminated in times of cutbacks. Management and old-line employees often still reject or suspect the value of molecularly based modeling until their own company or competitors score substantial successes. When they do buy in, there can be unrealistically high expectations, too.

Like any other tool, though, when people find a tool or skill or service that can make a difference, they seek it out. New graduates also create internal demand for specialists and for the tools themselves. Increasingly, they are familiar with the usefulness of these methods. They expect that these capabilities will be available as a matter of routine, just as they expect to have gas chromatographs or other analytical equipment.

Variations occur between and within industrial sectors:

- In *pharmaceuticals* and *agricultural chemicals*, all major companies now appear to include molecular modeling as a given necessity for identifying leads and analyzing binding sites. Modeling specialists are routinely on the product development teams.
- There are small, successful drug discovery companies like Pharmacopeia (Princeton NJ), Millennium Pharmaceuticals (Cambridge MA), and Vertex Pharmaceuticals (Cambridge MA) whose businesses are centrally built around a close interplay between modeling and experiment.
- However, most of "pharma" is not conscious of modeling value for their separations and other process or analytical needs. This appears to be due to pre-occupation with modeling's capabilities for product development, overlooking its potential in other aspects of the companies' needs. There has been some activity in these areas, for example at DuPont Pharmaceuticals.
- In the *fuels* and *chemical process industries*, including *polymers*, most companies have one to a few specialists. Some of these modeling groups have been created only recently and are thus on somewhat probationary status. Many others have established their value and are stable or growing.
- Large groups of specialists clearly imply that management sees modeling as a necessity. For example, the largest groups appear to be at Unilever and Sumitomo Chemical, each of which has 25 to 30 specialists.
- Likewise, the elimination or absence of modeling specialists in a company clearly implies that its management sees modeling as an unnecessary activity for the company's economic future. That may or may not be an absolute valuation by the company. In the case of Eastman Chemical, most research and development was eliminated in November 1999, including the two molecular-modeling specialists. By

contrast, when Amoco dramatically reduced its molecular modeling staff in the early 1990s, the decision appears to have been linked to dissatisfaction with their progress in modeling heterogeneous catalysis.

- In the *hard condensed-matter industries* (like *electronic materials* and *structural materials*), companies like Motorola, NEC, Toshiba, Fujitsu, and Corning have significant activity. This is a busy area of research. In most cases, it is a commitment to the future, while such work has established its value in a few companies. Process thermochemistry and kinetics are some of the properties being calculated, but most materials modeling of the materials themselves still uses continuum modeling with atomistic models. Band-gap theory is a notable exception to the latter.

(4) The principal role of molecular and materials modeling in the chemical industry is to speed product development and guide experiment.

For the discovery stage, this is most commonly and effectively done in the following ways:

- Developing mechanistic hypotheses and a rational, strategic approach to problem solving
- Providing understanding and design rules
- Stimulating new ways of thinking
- Ranking properties for molecules, materials, and formulations (best if in a QSPR model which can be queried)
- Aiming for predictive results, rather than just descriptive results
- Eliminating of dead ends—invaluable!
- Maximizing utility of data
- Broadening patents

For the manufacturing and process development stages, molecularly based modeling plays a key role in helping deliver accurate themochemical and kinetic data and in improving the yield, selectivity, or waste reduction for a given process.

(5) Science and methodology have made great strides, and successful applications encourage more.

- Quantum-chemistry accuracy is coming to large molecules by virtue of faster computers with more storage, new methodologies like hybrid methods and gradient-corrected DFT functionals, and sophisticated correlation-corrected quantum chemistry like G3, CBS, and ONIOM.
- Development and use of *ab initio* molecular dynamics enables better modeling of catalysis and materials (e.g., Car-Parrinello molecular dynamics, the T.U. Vienna code VASP, Nørskov's DACAPO code, Japanese CAMM/CAMP consortium codes).
- There are rapid improvements in modeling solvation and nonideal mixing properties through combinations of electronic-structure calculations with continuum solvation models (e.g., Tomasi's PCM and subsequent models, the COSMO-RS method developed at Bayer and spun off to a start-up company, COSMOlogic).
- Theories and modeling for "mesoscale structures" in fluids and solids are advancing (e.g., Dissipative Particle Dynamics from Unilever, the MesoDyn project originally from BASF and Groningen).
- Obtaining accurate kinetics and rate constants remains a challenge.
- More accurate force fields are needed than are presently available in today's commercial codes.
- An intersection is evolving with QSAR/QSPR correlations, informatics, combinatorial chemistry, and high-throughput screening—not just in pharmaceuticals.
- Few computer codes are available for phase equilibrium. Almost all from academia, and almost none are vendor-supported codes.

(6) Molecular modeling is being used to complement and interpret analytical measurements.

A remarkable number of properties are now calculable. For a wide range of experimental analytical techniques, predictions have become powerful aids to interpretation. Examples include:

- Ideal-gas thermochemistry ($\bullet H^{\circ}_{f,298}$, S°_{298} and C_p°)—Quantum chemistry. In some cases, predictions now have less uncertainty than the calorimetric measurements, effectively replacing them.
- Spectra for gas phase and solution phase, infrared and Raman—Quantum chemistry. DFT calculations are the current methods of choice to help predict and assign frequencies.
- NMR shifts for molecule structure analysis—Quantum chemistry.
- Crystal structures from crystallography data—Monte Carlo simulations with Rietveld refinement.
- Initial successes have been achieved in predicting retention times for gas chromatography by molecular simulations, predicting circular dichroism for optimal isomers using quantum chemistry, and predicting ion fragmentation patterns for mass spectrometry using reaction theory.

(7) Striking similarities exist among needs of different industrial sectors.

Different industrial sectors share many common technical interests in methods and in applications. Modeling solvation is a common and unresolved problem for the chemical and pharmaceutical areas. Modeling real mixtures is an issue for all sectors, whether in fluids, alloys, or at interfaces like grain boundaries. Likewise, the focus of drug and agricultural-chemical discovery is matching chemicals to the biochemical basis of a biological behavior, which will be the basis for predictive chemical toxicology. With predictive toxicology, chemical products could be designed that have a desired function but with less hazard than alternatives.

(8) The top needs required by the chemical industry can be summarized as "bigger, better, faster," more extensive validation, and multiscale techniques.

- "Bigger" means being able to do larger systems with greater complexity. Two examples are multi-reference wavefunction quantum mechanical methods involving complex transition-metal species and calculations of mesoscale behavior in multi-phase and composite systems. "Better" reflects the need for greater accuracy—e.g., activation energies, thermochemistry in condensed phases, and weak van der Waals interactions. "Faster" enables the simulation of rare-event processes such as predicting the thermal and oxidative stability of a plastic under a wide variety of weathering and the effect of a "dirt" particle on the fracture mechanics of a polymeric / inorganic filler composite.
- More extensive validation is required because of the complexity and range of modeling and simulation techniques required to solve real-world industrial problems. A modeling expert trained in one field needs to have extensive validation studies available to evaluate what level of theory is appropriate to model a given system, as well as to understand what the underlying assumptions and limitations are when using a new method in another field.
- Multiscale modeling is at the heart of technological or engineering modeling. Modeling methods that are essential at one scale may be useless at another. For example, all chemical reactions happen at a molecular level and require information from quantum chemistry, while neither force-field-based molecular simulations nor finite-element models are suitable. On the other hand, a dynamic physical process like conductive heat transfer can use molecular simulations effectively, convective heat transfer requires continuum-scale physics, and designing a bridge to carry a certain load in summer or winter requires a combination of engineering statics and materials properties which may only be measured. As soon as critical material properties are involved, property measurement is vital, yet understanding and extrapolation of those properties are tying ever more effectively back to the molecular level.

(9) Looking ahead, molecularly based modeling is the key to modeling a vital technology for the future: applying nanostructure and nano-scale devices.

Nanotechology is the technology of molecular-scale materials and processes. A typical atom has a van der Waals diameter of a few tenths of a nanometer; molecules and macromolecules are nanometer-sized and larger. All reactions and many macroscopic properties have their origins at this scale, linked to the

macroscale by statistical mechanics and solid-state physics. At the same time, the properties may be dominated by electronic structure, non-bonding interactions, or meso-scale, supermolecular behaviors. Each of these domains is the province of some type of molecular modeling, and it is increasingly possible to couple these domains to each other and the macroscale.

FINDINGS: INFRASTRUCTURE AND THE PROCESS OF APPLICATION

(1) The greatest economic value of molecular modeling is often not to create new substances and processes by itself, but rather to steer experimental development.

Give-and-take between modelers and experimentalists early in the development process has proven potent. In company after company, specialists argued that it is unrealistic to expect that molecular modeling would generate new catalysts or polymers, as unrealistic as expecting the same thing from NMR. At the same time, use in development teams has yielded big payoffs. At Dow, an internal meeting was held on developing homogeneous polymerization catalysts. To address questions and proposals made during the meeting, an impromptu session was held around an SGI workstation that had been quickly brought in, running simpler calculations with Spartan software and 3-D glasses. Successful patent filings for new catalysts came out of the meeting.

(2) Eliminating dead ends can be especially valuable.

In pharmaceutical discovery, modeling guides development by focusing attention in productive directions. A recent use is to help construct chemical libraries for combinatorial high-throughput screening, more promising subsets of the almost unlimited range of chemicals that might be examined. Likewise, a polymers example from Rhône-Poulenc was termination of an experimental elastomers development program, based on calculations that the approach could not succeed. Experimentalists were re-directed to more productive projects, and modelers were given internal credit for significant savings in that fiscal year.

(3) Failure to couple modeling sufficiently with experiments can lead to frustration, delays, and irrelevance.

In a case at Eastman Chemicals, a new modeler was assigned to get a mechanism for acid-catalyzed acyl transfer in alcoholysis of carboxylic anhydride. This was assumed to be the rate-limiting step for an ongoing process, and faster production was desired. No suitable transition state could be found. By probing more closely, the modeler found that operators added enough base that the reactant mixture was basic. A transition state was then found quickly. It was fast enough that equilibrium solubility instead proved to be limiting. It was then addressed separately.

(4) A healthy division has developed among specialists, non-specialist users, and internal clients.

Companies as different as Asahi Chemicals, Air Products, and Merck have initially tried to build a large user base for modeling. Eventually, they all decided that complicated or forefront problems benefit from specialists' expertise, but that the general user benefits most from low-level calculations and visualization. Non-users often make excellent use of calculated results, needing enough knowledge and motivation to contact and work with modelers but not doing the modeling themselves.

(5) Using molecular modeling for intellectual property extends beyond discovery.

There are numerous examples of patent positions and proprietary knowledge developed in whole or primarily with molecular and materials modeling. However, it is also used as follows:

- To extend initial discoveries, often broadening basic coverage by testing structural variations. One company modeled a small set of industry surfactants, then correlated the key property with a QSPR, tested and verified the result with a larger set, and ultimately discovered a new class of molecules for the job. Developing homogeneous metallocene catalysts is another good example.

- To defend intellectual property. Rohm and Haas successfully defended its patent for a urethane thickener by proving that the other company could not avoid making the Rohm and Haas thickener in its process.
- To build offensive patents, where a foothold of intellectual property is established in the market of a competitor.

(6) Computing platforms are shifting toward commodity processors, even for much high-performance computing.

Because of the focus on high accuracy for the biggest possible models, supercomputers have played an important role in past modeling. Since the late 1980s, workstations have been valuable for low-level calculations and visualization.

Workstations and Beowulf clusters, increasingly based on commodity processors, now are taking over even the largest tasks. Powerful single- and dual-processor PCs now achieve computational speeds faster than most previous workstations, and disk and memory costs are much lower. They are taking over many routine calculations. For large problems that benefit from parallel computation, the Beowulf approach is rapidly being adopted by companies worldwide. Multiple Intel Pentium or Compaq Alpha processors are coupled by a fast Ethernet switch and often powered by the Linux operating system, giving impressive parallel-computing performance for much less expense than a dedicated parallel supercomputer. Japanese supercomputer manufacturers noted that there will still be a significant market for ultra-powerful architectures, including supercomputer clusters. However, quantum chemistry codes do not necessarily parallelize well, and not all parallel problems are suited to this architecture.

(7) Software development must avoid becoming narrowly focused on the drug discovery problem.

These methods give extra intellectual power to the study and exploitation of life sciences. The visionary possibilities of precision and customized drugs, combined with the high profit margins presently available for successes in pharmaceuticals, have been powerful stimuli for development of new modeling and information technologies. At the same time, this has driven a consolidation of software companies and a relative neglect of software for other sectors like chemicals and materials. These sectors show growing demand for use of molecularly based modeling and for better software with more functionality. One consequence is that new software companies are beginning to appear again to fill this need.

(8) Technology-oriented modelers and modeling-savvy technologists are needed.

Many companies have begun using molecular modeling with a theoretical chemist and a technologist (chemical engineer or industrial chemist) working together. The mix of expertise is still valuable.

Over time, it has gotten easier but is still challenging to find theoretical chemists or physicists for the industrial environment. Likewise, new non-theoretical chemists are now better educated about these methods. Knowledgeable chemical engineers are especially needed. They are often well prepared in molecular thermodynamics, but relatively few have experience with molecular simulations and computational quantum chemistry.

CHALLENGES FOR THE SCIENCE AND APPLICATION OF MOLECULAR MODELING

For future successes, both science and its implementation are important. It is not only necessary to identify or develop the right science for a given application, but new tools and suitable infrastructure must be built within companies and scientific communities. Specific topics can be cited:

1. *Algorithms.* Improved computational algorithms are needed to cope with larger molecules or sets of atoms. As we can tackle bigger problems, we invariably need to tackle even bigger ones. One measure of the problem size is the number of atoms, N. Scaling of problem size by N^3 to N^7 quickly limits the feasible targets of *ab initio* calculations, and a molecular-dynamics simulation of 1000 atoms for a second might as well be infinity when computed in femtosecond time steps. More degrees of freedom

mean more complicated optimizations, as in protein configuration analysis, self-consistent-field solutions, or transition-state searches.

Introducing new science may necessary in much the same way that electronic density functional theory was adapted from its base in solid-state physics. Recent advances in solvation, one of the central problems for technological applications, have come from redefining the reference condition from ideal mixtures of non-interacting molecules to a perfectly screened molecules in perfect conductors. However, new algorithms can drastically affect solution of the basic quantum and statistical mechanics. Consider the QM/MM methods.

2. *Exploiting parallel computation.* Parallelization has great promise for large systems, yet conventional approaches to quantum chemistry have few parallelizable features to exploit compared to molecular simulations. Methods and software are needed to parallelize problems and to develop parallel code more easily. For example, Lester and co-workers have shown that a Quantum Monte Carlo method can be used to evaluate the integrals used to solve the Schroedinger equation with great accuracy. On single-cpu computers, the method is quite slow, but with hyper-parallel computers, it could deliver fast, accurate calculations even for quantum chemistry calculations on large molecules.

   Likewise, parallelizable problems are still limited by memory usage or message passing. Dropping the dimension of the problem from N^3 to N^1 shrinks the problem a million-fold for a problem involving 10 oligomers of 100 atoms each. Advances like the "Order(N)" methods could have great impact.

3. *Automating the choice as well as the methods.* A different issue is automation of method choice. For example, quantum chemistry calculations may be carried out in a hierarchy of increasing accuracy (Pople's "model chemistry" approach) and cost, and molecular simulations may be carried out with force fields of increasing detail and accuracy—and cost. In both cases, there is a trade-off. BASF developed and tested a semi-automatic selection algorithm based on the desired property and accuracy (CrunchServer), finding requested information in the database or delivering the request for human assessment of choices. It was discontinued because of unresolved challenges, but it foreshadows information management methods for the future.

4. *Computing hardware and operating systems.* Faster CPUs, more memory, and parallel configurations offer new power and new challenges. Different approaches to molecular modeling place different demands on hardware and operating systems. So do the varying needs of users. Base-level platforms must be integrated with the computing and visualization necessary to solve smaller but important problems effectively. User interfaces must be transparent windows on the task, methods, and results, despite use on diverse, constantly evolving platforms. Similarly, it is increasingly desirable to carry out code generation, translation, and documentation using cross-platform visual programming tools and techniques like literate programming.

5. *Data management.* As computational power increases, the desire to probe the real complexities also demands more sophisticated management and exploitation of data, both computed and measured. While some problems may reduce to a single key result (e.g., a heat of formation), most also have a large amount of accompanying detail. Polymer modeling is a good example. Consider the normal-mode analysis of frequencies for 100,000 MW polyethylene ($C_{7143}H_{14288}$). For any configuration, each of the 21,431 atoms has 3 time-dependent position coordinates, simplified to 3 coordinates and 3 components of an oscillation motion vector. It gets worse. These motions may be nonharmonic, the polymer may be highly branched, the polymer will exist in some molecular-weight distribution, and analysis of a single chain will be inadequate to capture the amorphous or crystalline morphology of the material. Furthermore, the interest may be in small-molecule diffusion in a polymer, a polymer melt, homopolymers of more complicated monomers than ethylene, block or random co-polymers, polymers with plasticizers and other additives, solvated polymers, stereochemistry, radical or ionic or condensation or metal-catalyzed polymerization, or polymer degradation. Biopolymers add another layer, as do comparisons with noisy data. The task seems impossible, yet even now polymer modeling is powerful, and as complex a behavior as protein folding can be attacked by crucial simplifications.

6. *The experiment-modeling interface.* Pure correlations like QSPR and QSAR (quantitative structure-property or structure-activity relations) already allow prediction and interpretation of practical properties like toxicity or octanol-water partition coefficients. Prediction of measurable properties is crucial, but so

is measurement of predictable properties. A challenge for molecular modeling is to keep up with the fast and effective approach of combinatorial chemistry, which in principle is quite Edisonian. Already, computational chemistry and information technologies have become important for systematizing the array of possibilities to be tested experimentally (e.g., see site report on Pharmacopeia, Inc.). Combinatorial chemistry can be a powerful way to avoid the limitations of imagination, but the best combinatorial chemistry also builds on relevant chemical principles. Computations need to aid the design and interpretation of such experiments, leveraging the appropriate time/accuracy balance of different theoretical approaches.

7. *Problem analysis.* Successes in industrial application have rested on identifying the crucial issues or questions. The Eastman Chemicals acylation problem, noted above, is a good example of the need for modelers and non-modelers to analyze the problems together. Understanding what these methods can and cannot do is crucial to successful problem analysis.

8. *Personnel infrastructure.* Two personnel issues have been crucial to initial successes in industrial application of computational chemistry: (1) identifying people who know or are willing to learn both the computational chemistry and the applications, and (2) having advocates among management or clients who recognize the appropriate uses of the methods. The largest industrial groups of specialist molecular modelers consist of 25-30 people, and most are much smaller in size. In the chemical process industries, it is common to find only one or two computational chemists in a company, frequently paired with a more technologically oriented chemist or chemical engineer. At the outset, they educate each other. As time proceeds, they educate others in the possibilities and limitations of the computational tools, most fruitfully by aggressive participation in development teams at the earliest stages. In contrast, use of technical service providers and consultants has had limited success, but has sometimes led to elimination of the activity when internal clients failed to take advantage of tools they didn't know about or didn't trust.

9. *Validation.* Method validation is often an unstated but crucial need. Comparisons by individual researchers and by software vendors like Gaussian, Inc. are useful for general guidelines. Other, more comprehensive comparisons will be valuable. For example, NIST is building a Computational Chemistry Comparison and Benchmark Database (http://srdata.nist.gov/cccbdb/) and a Sicklist Database of known problems for *ab initio* methods (http://srdata.nist.gov/sicklist/). As a multinational effort, IUPAC has a task group for "Selected Free Radicals and Critical Intermediates: Thermodynamic Properties from Theory and Experiment," led by Dr. Tibor Berces of the Hungarian Academy of Sciences' Chemical Research Center.

10. *Credibility.* A final key need is development of substantive credibility. Successes within an organization are the ultimate criteria, but awareness of outside successes has proved to be a powerful spur to using these methods. Two dangers have been the desire for overwhelming successes and the strong impact of attractive visuals. There are cases of computational chemistry groups being eliminated because of undue expectations from management and colleagues. These expectations had been heightened by promises to succeed at strictly long-term tasks like *de novo* catalyst design, reinforced by realistic-looking visual images. In contrast, Dow researchers note that their successes in modeling polymer properties were made possible in part by their shorter-term successes in calculating needed ideal-gas thermochemistry.

VISION

There are two kinds of speculation about the future: Things we reasonably expect ("near-term") and things we only dare imagine ("long-term"). For example, in 1970 it was reasonable to expect computing would be more important by the beginning of the twenty-first century. On the other hand, it would have been a luxurious vision to suppose that individuals might have their own computers. It would have been even more like science fiction to suppose that nontechnical businesspeople and schoolchildren alike could have their own powerful palm-sized computers with wireless connections to a spiderweb of computers, information, and people around the world. On the other hand, it was 1929 when Dirac pronounced that all the physical laws underlying chemistry were known—although their application was hindered by the exact equations being unsolvable.

For molecularly based modeling, advancing computer power drives the near-term promise:

- The value of computed results will be greatly enhanced when they can be generated near-instantly, but they become invaluable when their meaning is clearly linked to the application's or developer's need.
- Conventional methods of *ab initio* calculation will be much faster for a given problem, although they will still be unable to handle much bigger problems so long as high-accuracy methods increase in cost as a function of $N_{atoms}^{\ 6}$ or $N_{atoms}^{\ 7}$.
- Parallel supercomputing will be increasingly available, offering great advantages to molecular simulations that parallelize well.
- Experimentalists and management will not only become used to accepting the use of molecular modeling, but they will expect it.
- There is so much recognition of the need for validation and testing of methods that expert systems should become available to guide the novice or intermittent user in making the best calculations.
- Vast amounts of numerical data and qualitative information will demand more reliance on correlation, mining, and visualization of results.

Long-term, we can speculate about science, applications, and infrastructure separately, but the complicating factor is that present paradigms will be turned on their heads by the ubiquity of computing processors. Already, changes in computing power and architectures outpace the development of software that takes advantage of them. Raw power will make calculations of unthinkable size thinkable, but the real surprises lie in the unimagined. When computers can be printed or grown, it may become easier to design a new computer for a problem than to design software, eliminating the idea of operating systems. Stochastic methods will dominate if all of parameter space can be rapidly examined. Ironically, physicochemical data will become limiting; how can one confirm that a method's uncertainty is better than an experiment's if the experiment is the standard? At the same time, great advances are being demanded in knowing and controlling properties and behaviors of individual atoms and small groups of atoms, whether for nanostructured materials or gene manipulation or quantum computing. Understanding in turn demands models that are consistent with physical reality; they are the molecularly based models that are the subject of this analysis.

PERSPECTIVE

The United States leads this field in many scientific areas. However, Canada has particular strengths in DFT methods and homogeneous catalysis; Europe in heterogeneous catalysis, mesoscale, and materials modeling; and Japan in materials modeling and special-purpose computing. Major government-industry initiatives are underway in Europe and Japan, notably in multi-scale materials modeling and in development of chemistry-capable *ab-initio* molecular dynamics codes. In European and U.S. assessments of nanotechnology, it was also concluded that to advance the field most quickly—and competitively—the need is acute for applying new and existing methods of molecularly based modeling.

Company activities and challenges are generally the same around the world. Paradoxically, quantum mechanics and statistical mechanics are well-established fields, but their quantitative applications have been held back by the computer technology for using them. Success and failure always depend on the right people working on the right problem with the right tools, and activities in molecular modeling have been vulnerable. The process and organization of these activities have proven equally important to the science and technical implementation.

Raw computing power can't buy success, but advances in computing power, visualization, and information technology have made many new applications of molecular modeling feasible. Growth in parallel computation ability has not kept pace with advances in parallel computer hardware, largely because methods lag for construction and optimization of user codes. Ironically, using specialized parallel hardware (SGI Origins, for example) appears to be slowing as the economics of commodity CPUs drive rapid increases in their use (e.g., in Beowulf clusters). User interfaces with software remain a problem, especially for generation of thermochemical and kinetics parameters.

Advances in practical applications during the 1990s were remarkable, though. The head of corporate research at one chemical company in 1999 went so far as to say that it was clear that the company's success in applying molecular modeling would be a central basis of its success ten years from now. Not all management feels so strongly, but the present widespread penetration of these methods speaks for their perceived and demonstrated importance.

In education and technical communications, both for students and for practicing scientists and technologists, knowledge of molecularly based phenomena and modeling is becoming imperative. Continuum models of materials and of macroscopic behaviors will continue to be important. However, all reaction chemistry occurs at the molecular scale, molecular interactions are the basis of continuum properties, and life sciences are becoming ever more clearly linked to their chemical underpinnings. Molecular theories and modeling, including electronic-structure theories and modeling, are likely to become the *lingua franca* for advances in much of science and engineering.

CHAPTER 1

INTRODUCTION

Phillip R. Westmoreland

GOALS AND GENERAL APPROACH

The goals of this analysis were (1) to assess state of the art in molecular modeling, from quantum chemistry to empirical molecular simulation of macroscopic systems, and (2) to assess how effectively and extensively theoretical advances are being transferred to applications in the chemicals, life sciences, and materials industries. A small committee of experts with diverse expertise was chosen to gather information from site visits, interviews and the literature:

- Phillip R. Westmoreland, Chemical Engineering, University of Massachusetts, Amherst (Chair)
  - Measurement and quantum-chemical modeling of reaction kinetics for combustion and polymer development
- Peter A. Kollman, Pharmaceutical Chemistry, University of San Francisco (Vice Chair)
  - Development and application of computational approaches to understand structure, energies and reactions of organic and biochemical molecules, for example the AMBER molecular simulations code and force-fields
- Anne M. Chaka, Lubrizol Inc. (at NIST from May 2001)
  - Industrial computational chemistry and physics, notably applied to problems of lubricants and corrosion
- Peter T. Cummings, Chemical Engineering, University of Tennessee and Oak Ridge National Laboratory
  - Methods for predicting physical properties in systems of industrial interest such as supercritical aqueous solutions, alkane fluids, and polymer solutions using molecular simulations and massively parallel processing
- Keiji Morokuma, Chemistry, Emory University
  - Methods and applications of computational quantum chemistry, including the IMOMO and ONIOM methods and their applications in combustion chemistry and homogeneous catalysis
- Matthew Neurock, Chemical Engineering, University of Virginia
  - Modeling of heterogeneous catalysis using electronic density functional theory and kinetic Monte Carlo methods
- Ellen M. Stechel, Chemistry and Environmental Science Department, Ford Motor Company/Ford Research Laboratories
  - Industrial application of theoretical chemistry and physics, particularly of condensed-matter physics and catalysis, and perspective of industrial research management

- Priya Vashishta, Physics and Computer Science, Louisiana State University
  - Modeling materials behavior using atomistic simulations for very-large-scale problems, exploiting massively parallel supercomputing to study fracture, crack propagation, and nanostructures

These panel members were chosen and the study was sponsored by a diverse group of U.S. government research and development agencies, under the leadership of the National Science Foundation's Directorate for Engineering (in particular the Assistant Director's Office and the Division of Chemical and Transport Systems):

- National Science Foundation
  - Engineering Directorate
  - Mathematical and Physical Sciences Directorate (Division of Materials Research)
- Department of Energy
  - Basic Energy Sciences/Office of Science
  - Office of Industrial Technology
- Department of Defense
  - DARPA
  - Air Force Office of Scientific Research
- National Institutes of Health
- Department of Commerce (National Institute of Standards and Technology)

As a key part of the study, the panel made 53 site visits in Europe, Japan, and the United States, mostly to commercial organizations. The panel chose to rely on personal knowledge and the open literature for including other industrial, academic, and governmental activities. Detailed reports were prepared on activities at many of these organizations, reviewed by those organizations, approved by them, and included in Appendix B of this report. These and other activities were analyzed and evaluated. Related funding of these activites by the United States government was also analyzed and estimated to be about $200 million, as summarized in Appendix C. Finally, a glossary is provided in Appendix D, and an index of the mentioned companies is provided as Appendix E.

Representatives of the sponsoring agencies attended some of these visits and contributed site reports that are included in the appendices: Raul Miranda (NSF), Anna Tsao (IDA, representing DARPA), Karl K. Irikura (NIST), Sharon C. Glotzer (NIST), and Randall S. Jones (ITRI-WTEC).

Study management was handled by the WTEC Division of the International Technology Research Institute, based at Loyola College in Baltimore (http://itri.loyola.edu/; http://www.wtec.org/).

REASONS FOR THE INTERNATIONAL STUDY

During the 1990s, molecular modeling began to be a valuable tool of science and engineering. Accuracy, efficiency, computational power, and appreciation of these methods reached the point where they provided valid insights and allowed sound decision making.

Consequently, its use has spread into many companies, yet paradoxically, industrial use is still in its infancy. The methods have been used most by a small number of specialists, although new employees frequently expect modeling capabilities to be available. The most touted potential and successes have been in identification and design of pharmaceuticals, a high-profit application. Less well known is the impact occurring in design and development of homogeneous catalysts, where teamed modeling and experiment are leading to dramatic improvements. Still more unsung is the way that these approaches are transforming routine tasks of finding physical and chemical properties and trends, especially in process and product development.

Theoretical advances have come from around the world, but clusters of researchers in different regions have established different strengths. For example, the United States has demonstrated particular strengths in pharmaceuticals, force-field development, phase-equilibrium modeling derived from classical molecular thermodynamics, and homogeneous catalysis. Canadian researchers are also deeply involved in homogeneous catalysis and electronic density-functional theory. Among the areas of strength in Europe are kinetics of heterogeneous catalysts and mesoscale modeling. Achievements in Japan, India, Australia, and other countries are recognized but are not known as well in the United States.

A number of specific successes in industrial applications are known, but technical details have been revealed for relatively few. Details can sometimes be published and presented when there are modest advances in fundamentals, industrial successes in interesting areas that nevertheless prove not to be commercially promising, and even relatively small failures. In contrast, technical details of major successes can rarely be discussed, and major failures of development are left behind, undiscussed and sometimes left unanalyzed even internally.

Non-technical details can be extremely instructive—effectiveness of organizational structures, choice of hardware and software, selection and training of project personnel, approaches for effective communications, and so on. These topics are usually not suitable for technical journals, so they tend to be passed on (if at all) by informal discussion and anecdotes.

Several earlier studies have been sponsored by U.S. government agencies to review different aspects of applying molecular and materials modeling, including the following:

- *Mathematical Challenges from Theoretical/Computational Chemistry* (http://books.nap.edu/html/mctcc/, 1995)
- NSF *Workshop on Future Directions in Molecular Modeling and Simulation: Fundamentals and Applications* (http://flory.engr.utk.edu/nsf.html/, 1998)
- DOE/OIT *Technology Roadmap for Computational Chemistry*, Workshop held by DOE/OIT on March 16-17, 1998 (http://www.ccrhq.org/vision/index/roadmaps/CompChem%20Roadmap%2092599.pdf, 1999)
- DOE *White Paper on Computational Materials Science and the Strategic Simulation Initiative* (http://cmpweb.ameslab.gov/cmsn/otherlinks/executive.html, 1999)
- DOE *White Paper on Computational Structural Biology and the Strategic Simulation Initiative* (http://cbcg.lbl.gov/ssi-csb/Program.html, 1999)
- *Impact of Advances in Computing and Communications Technologies on Chemical Science and Technology: Report of a Workshop*, Chemical Sciences Roundtable, National Research Council, National Academy Press, 1999 (http://books.nap.edu/html/comp_chem_sci/)

The analysis described here builds on these earlier studies, aiming to encompass a still wider area of topics and geography.

SCIENTIFIC AND TECHNOLOGICAL SCOPE

The scope of this analysis may be classified by scientific, technological, and infrastructure issues. Differences in regional strengths, differences in governmental roles, and differences among industry, academic, and governmental interests are also informative.

In terms of molecular science, "molecular simulations" and "computational quantum chemistry" are the two main thrusts, coupled with modeling and interpretation of the results ("informatics"). Molecular simulation is based on classical Newtonian physics, modeling interactions within or between molecules using correlations of the interaction forces (force fields). Systems are modeled either deterministically (molecular dynamics, which integrates classical equations of motion) or as stochastically varied cases (Monte Carlo methods). In contrast, computational quantum chemistry is based on quantum physics, primarily applied to

the electronic structure of atoms or molecules. The immediate results are wavefunctions or probability density functionals describing electron states.

The distinction is pragmatic. In molecular simulations, simpler models with many parameters make it possible to model large collections of atoms and molecules. Quantum chemistry can provide greater accuracy but is restricted to smaller molecular size by its complexity and cost. Hybrids of the two methods offer new possibilities, along with limitations.

As a result, molecular simulations are used to model ensemble properties and behaviors. Examples include P-V-T relations, phase equilibrium, transport properties, structures of synthetic and biological macromolecules, and docking of one molecule against another.

Quantum chemistry is essential when chemical bonds are made or broken, but it is also used when force parameters are unknown or not applicable. High-level wavefunction methods with large atomic-orbital basis sets currently remain the standard, but density functional methods are used increasingly and with increasing accuracy. Semi-empirical methods are another important subset, which although quite parameterized, capture often-useful amounts of electronic structure in faster, more affordable computations. Some general areas of utility are force-field development and calculation of thermochemistry, kinetics, optical properties, and NMR shifts.

The entire field may be summed up as "computational chemistry" or "molecular modeling," but some researchers in the area find these terms distastefully inaccurate. To many people, the term "computational chemistry" automatically implies computational quantum chemistry. On the other hand, this term may be helpful when used more inclusively, bringing in the use of reaction theories like RRKM, master equation theory, or variational transition state theory to compute kinetics. Similarly, it is a fair criticism to recognize that much of molecular simulations is "materials modeling" by classical physics, not modeling of molecules or of chemistry.

Unresolved theoretical issues range from simple to conceptually rich. For example, the seemingly simple problem of properly interpreting internal rotations is a serious problem for prediction of thermochemistry and kinetics, complicated further when in a condensed phase or in solution. Simulation of polymer distortions in flows relies on lumping multiple atoms, collapsing electronic structure and sensitivity to interatomic forces into pseudo-atoms. As a third example, the assignment of poly-gaussian hydrogenic orbitals to heavy elements breaks down because of difficulty in modeling more complicated distortions in valence electrons, active participation of numerous non-ground-state electronic configurations, and relativistic effects on electrons near the massive charge of the nucleus.

SCOPE BEYOND THE SCIENCE OF MOLECULAR MODELING

To achieve successes, it is important not only to identify the correct existing science for a given application, but also to build new tools and to build useful infrastructure within companies and scientific communities. Some of the topics that need consideration are:

1. *Algorithms.* Improved computational algorithms are needed to cope with larger molecules or sets of atoms. Because interactions among atoms are the heart of the matter, solution time and storage requirements typically rise faster than the number of atoms, N. Methods like "Order(N)" methods could have great impact.

2. *Exploiting parallel computation.* By dividing problems among separate cpu's, parallel computing offers the possibility of proportionately speeding up solutions. However, reaching this goal requires problems (or parts of problems) that can be divided equally, algorithms to take advantage of parallel structures, and ways to write and debug parallel codes efficiently. Many quantum-chemistry codes parallelize poorly, partly due to the methods they employ. An example alternative is the Quantum Monte Carlo method of Lester and co-workers.

3. *Automating the choice as well as the methods.* One role of the modeler is to choose the best method for a particular problem with particular accuracy requirements. Such decisions might be automated to the

benefit of both more and less-skilled practitioners. Careful comparisons of different methods, different computing configurations, and different objectives are necessary.

4. *Computing hardware and operating systems.* Different approaches to molecular modeling place different demands on hardware and operating systems, as do the varying needs of users. The mix of supercomputers, personal computers / workstations, and distributed computing is another issue. Integration of platforms, visualization, user interfaces, and cross-platform visual programming tools are required.

5. *Data management.* Not only the computational needs but also the amount of results go up dramatically with problem size. This challenge demands sophisticated management and exploitation of both experimental and computed data. To be useful, the huge amount of new information must be managed in such a way that it can be turned into empirical and theory-based relationships.

6. *The experiment-modeling interface.* Correlations like QSPR and QSAR (quantitative structure-property or structure-activity relations) relate observable properties to calculated molecular-scale properties. These correlations may be purely empirical, but they are normally successful when the calculated properties are well chosen. Likewise, the measured properties should be chosen both to reflect performance and to be based on physical and chemical principles. Prediction of measurable properties is crucial, but so is measurement of predictable properties. Combinatorial chemistry is a good example, where the greatest success comes by building libraries of candidates using sound chemical principles.

7. *Problem analysis.* Successes in industrial application have rested on identifying the crucial issues or questions. A common experience is that described by Peter Margl when he was at Eastman Chemical, responding to a request from the process development group to identify a mechanism and rate of acylation. That group had made the specific request because they assumed this step was crucial to an anhydride alcoholysis. After working through to the answer to the requested problem, the computational chemists returned to the process development group and uncovered the real goal—faster alcoholysis—and quickly identified equilibrium solubilities of the reactants as rate-limiting. Also, because of misunderstanding about what these methods can and cannot do, collaborative development of the tasks is crucial to successful problem analysis.

8. *Personnel infrastructure.* A key personnel issue is having the right technical people, who individually or as colleagues know both the modeling and the technological applications. It is just as important to have management and clients who accept the usefulness—and limits—of the methods. Both sets of people need adequate education and experience.

9. *Validation.* Validation usually refers to testing of models against data to establish their validity or limits of validity. This task may be understated or neglected, but it is needed crucially.

10. *Credibility.* Credibility has two aspects: establishing it and preventing its loss. Credibility must be built honestly by achievement. Failures, properly understood, do not necessarily damage credibility. Rather, the worst threat has repeatedly proven to be allowing expectations that are unrealistically high. Ironically, computer visualization has been a key to rapid communication of results, but it has also been destructive to modelers' credibility because an attractive picture can be overly persuasive. The inevitable result is heightened skepticism. One must work on suitable problems using the right tools.

ORGANIZATION OF THE ANALYSIS

The chapter structure of this report is in the order of science, application, infrastructure issues, and international comparison. Chapters 2, 3, and 4 are science-focused, examining the status of computational quantum chemistry (including hybrid methods, kinetics, and thermochemistry), molecular simulations, and informatics. Chapters 5, 6, 7, and 8 examine technological applications: the chemical processing industries, applied life sciences, physical and electronic materials, and catalysis. Chapter 9 presents a range of infrastructure and management issues important for successful application of molecularly based modeling. Chapter 10 is a global comparison of industrial, academic, and governmental activities. Finally, Chapter 11 offers a review of the study's findings and a vision of its implications.

Supporting materials in the appendices are a valuable part of this work. Following biographies of the panelists, an extensive set of site reports is presented, which provides an important basis for the analyses. Finally, a glossary of terms and acronyms is included.

CHAPTER 2

SCIENCE: ELECTRONIC STRUCTURE, THERMOCHEMISTRY AND KINETICS

Keiji Morokuma

INTRODUCTION

Electronic structure theory based on quantum mechanics is one of the most fundamental tools for molecular and material modeling. Quantum mechanics enables modelers to calculate the structure, energy and properties of a molecule or an assembly of molecules. Using this energy, one can calculate the structure of the reactant and the product, as well as that of the transition state that connects the reactant to the product in the chemical reaction. Thermal properties of molecules (thermochemistry) and the rate of reactions (kinetics) can also be calculated using this energy. From the quantum mechanical calculations, one can also obtain spectroscopic properties of molecules (UV, IR, Raman, NMR, NQR, Mössbauer, etc.).

In electronic structure theory, given the position of atomic nuclei, $\boldsymbol{R}$ (under the Born-Oppenheimer approximation), the Schrödinger equation for motion of electrons $\boldsymbol{r}$:

$$H(\boldsymbol{r},\boldsymbol{R})\,\Psi(\boldsymbol{r};\boldsymbol{R}) = E(\boldsymbol{R})\,\Psi(\boldsymbol{r};\boldsymbol{R}) \qquad (2.1)$$

is solved. Here, $E(\boldsymbol{R})$ as a function of the nuclear coordinates is the potential energy surface (PES), which determines the equilibrium and transition state structures of a molecule as well as vibrational normal coordinates and frequencies. $\Psi(\boldsymbol{r};\boldsymbol{R})$ is the wave function that determines electron density and various properties, such as dipole moment and electrostatic potential. Various approximate methods for solving the Schrödinger equation are available for molecular modeling, semi-empirical molecular orbital (MO) methods, *ab initio* MO methods, and density functional theory (DFT) methods.

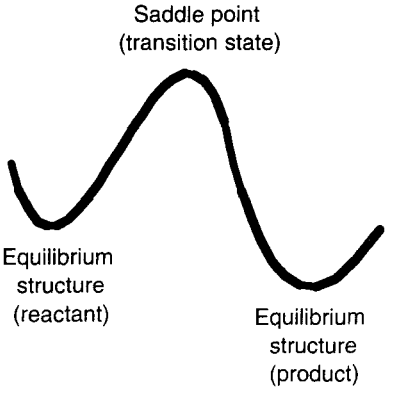


Fig. 2.1. Stationary points on a potential-energy surface.

SEMI-EMPIRICAL MO METHODS

Semi-empirical MO methods neglect most of the two-electron integrals in solving the Schrödinger equation and use experimental results (parameters) to adjust integrals to obtain good results at very low cost. Some of the methods still used in modeling, especially for large biological molecules, are MNDO/d (Thiel 1996), AM1 (Dewar et al. 1985) and PM3 (Stewart 1989). The INDO/S (ZINDO) (Zerner 1995) has been used extensively in industrial applications. This method, with parameters readjusted especially for absorption spectra, is a powerful tool for excited states of large molecular systems and molecules containing transition metals. Examples of applications include electronic spectra of $C_{60}O_2$ and $C_{60}O_3$ and those of $[Fe_2(\mu\text{-}S_2)(P(o\text{-}C_6H_4S)_3)_2]^{2-}$. This inexpensive method has been used to calculate nonlinear optical properties of large molecules, in particular in industrial applications (e.g., see site report in Appendix B for Mitsubishi).

Some of the recent developments for semi-empirical methods are intended for applications to even larger molecules. The MOZYME method (Stewart and Stewart 1999) performs direct calculation of electron density and energy without diagonalizing the semi-empirical SCF (self-consistent field) matrix and can be used for geometry optimization of large molecules. Applications so far include polypeptides consisting of 248 amino acid residues. The Divide-and-Conquer method (Dixon 1996) originally developed for the density functional theory, divides a large molecule into smaller segments, performs semi-empirical SCF calculations for segments, and obtains the electron density of the entire molecule "adding" the density of segments, which is used to calculate the energy of the entire system. Geometry optimization can also be carried out. One application is in modeling lysozyme, consisting of 1960 atoms in 140 residues. Another new development of semi-empirical methods, SCC-DFT (Frauenheim et al. 2000), based on the density functional theory, takes into account self-consistent redistribution of charges and accomplishes a substantial improvement in energetics over AM1/PM3. This method is being applied to biomolecular systems and large molecular materials. These new methods are likely to be used in the near future for industrial applications.

Summarizing semi-empirical MO methods, their largest merit definitely is low cost. Despite very significant speed-up relative to *ab initio* methods, they are still much faster than any *ab initio* method. They are more expensive than the molecular mechanics (MM) method, but they allow breaking of bonds and take electronic effects explicitly into account, which MM cannot. Important shortcomings of semi-empirical methods are low reliability (qualitative at best, and particularly poor for transition states) and lack of reliable parameters for transition metals (therefore not applicable to most homogeneous and heterogeneous catalysis modeling). For the near future, the semi-empirical approach will remain a quick and inexpensive alternative to *ab initio* calculations in industrial and biomolecular applications.

AB INITIO METHODS

In the *ab initio* MO methods, the Schrödinger equation is solved "from the beginning" (which is not "from first principles" as *ab initio* is frequently translated). They usually express the molecular orbitals (MO) as linear combinations of finite number of basis functions:

$$\varphi_i = \sum_{\nu=1}^{Nbasis} C_{i\nu}\chi_\nu \tag{2.2}$$

The basis set can start as small as "minimal" (one basis function per valence shell) and can be improved systematically by going to "double zeta" (two per valence shell), to "double zeta plus polarization", to "triple zeta plus double polarization plus higher angular momentum plus diffuse functions," eventually approaching the complete set. The total electronic wave function is expressed as a linear combination of Slater determinants (SDs):

$$\Psi = \sum_I D_I \Phi_I, \Phi_I = (N!)^{-1/2} \sum_P P(-1)^P \{\varphi_1 s_1(1)\, \varphi_2 s_2(2)\, \varphi_3 s_3(3) \cdots \varphi_N s_N(N)\} \tag{2.3}$$

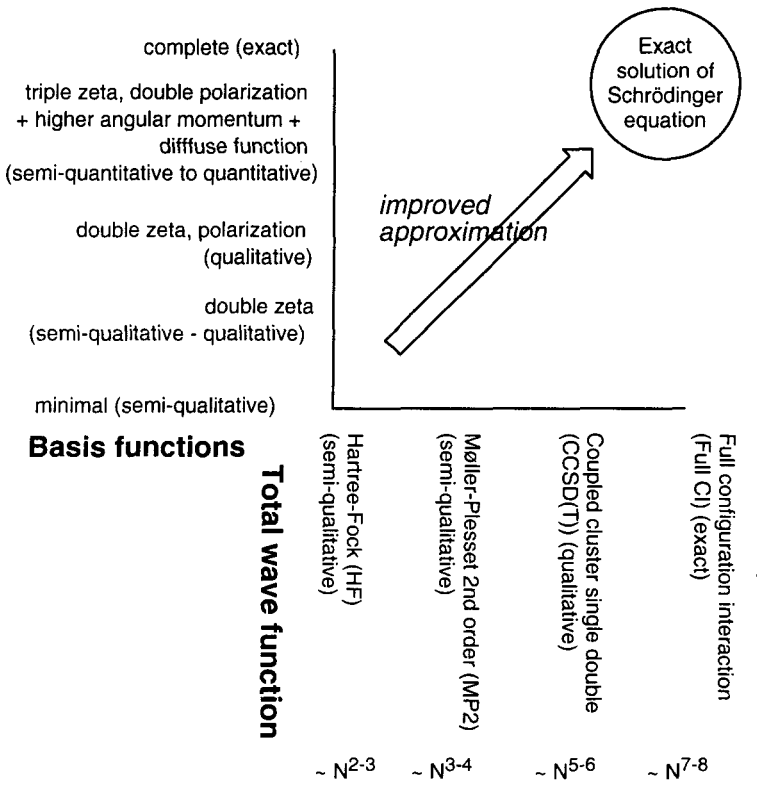


Fig. 2.2. Hierarchy of methods and basis functions to approach a full, exact solution to the Schrödinger equation.

The simplest form of the total wave function is a single SD, which is called the Hartree-Fock (HF) approximation. To include "electron correlation," one has to use more than one SD. Extensive studies in the last 20 years have established that, with an increasing order of sophistication and accuracy, Møller-Plesset 2nd order perturbation (MP2) and coupled cluster single and double with perturbative triple excitation (CCSD(T)), among others, have been found to be reasonable approximations of a full configuration interaction (Full CI) calculation. The CCSD(T) calculation with a very large basis function, or an extrapolation scheme such as G2, G3 or CBS for this target, has been found to provide energetics (atomization, ionization, and others) within a few kcal/mol of average error from the experiment (Yarkony 1995).

One of the hottest topics in the *ab initio* method is the multi-reference coupled cluster (MR-CC) method. The standard CCSD(T) improves the approximation starting with the Hartree-Fock (single SD) wave function. Some molecules or states (such as diradicals or diradicaloids) require more than one reference configuration (SD) as the starting point; in theoretician's language, "non-dynamic electronic correlation as well as dynamic correlation has to be taken care of." Active development is taking place presently to perform coupled cluster calculations starting with the multi-configurational reference function, called MR-CC. In a few years, reliable codes will be available for general use, and the method is expected to be very accurate (quantitative) for difficult molecules but would be very expensive (Mahapatra, Datta, and Mukherjee 1997).

Another hot topic in *ab initio* methods is electron correlation calculation using localized MOs. In traditional methods for electron correlation, delocalized MOs are used as the basis. However, electron correlation is a short-range problem, and localized MOs (LMOs) often give better convergence. As the size of molecule increases, the number of important LMOs will increase only linearly with the size, and it may be possible to accomplish "linear scaling ($\sim N^1$)" in high-accuracy calculations. Some of the successful implementations include LMO-MP2 (Rauhut 1998), LMO-CI, and LMO-CC (Laidig, Purvis, and Bartlett 1982). These LMO-based methods may replace the traditional electron-correlation methods in the near future. These new high-

end methods in the *ab initio* MO method will be used in molecular modeling to improve the energetics needed in thermochemistry and chemical kinetics, which are of major concern in industrial applications.

Some important applications of the *ab initio* MO method are for electronic excited states, especially in connection with spectroscopy and photochemistry. Most excited states are multi-configurational in nature (cannot start with one SD), and often more than one excited state needs to be calculated at the same time. For this purpose, no unambiguous method (comparable to CCSD(T) for the ground state) has been established. Commonly used is the CASSCF (complete active space self-consistent-field) method, which determines MO coefficients C_{ir} and CI coefficient D_j at the same time and is considered to be a quite reasonable starting point:

$$\varphi_i = \sum_{\nu=1}^{Nbasis} C_{ir}\chi_r\,,\ \Psi = \sum_I D_I \Phi_I \tag{2.4}$$

The MR-PT2 method (Fulscher, Andersson, and Roos 1992) usually starting with CASSCF results, takes into account "dynamic correlation" using the Møller-Plesset-like 2nd-order perturbation. The MR-SDCI method usually starts also with CASSCF results and improves the energy of several excited states at the same time by performing CI calculations with single and double excitations. This method can be used not only near the equilibrium geometry but also far away from it, and it is suited for studies of chemical reactions. The state-of-the-art methods for excited states are SAC-CI (symmetry-adapted cluster expansion method, followed by CI) and nearly equivalent EOM-CC (equation of motion method based on the coupled cluster expansion) (Nakatsuji 1979; Watts, Gwaltney, and Bartlett 1996), starting with the highly correlated coupled cluster method (such as CCSD) for the ground state and performing CI calculation including single and double excitations. These methods have been found to give excellent results for excitation energies, which is ideal for electronic spectra. Performance away from the equilibrium geometry is yet to be fully tested.

The importance of conical intersections in photochemical reactions has now been recognized widely (Klessinger and Michl 1995). As shown in Fig. 2.3, the conical interaction brings the molecule in an upper state to the lower state though a cone in the potential energy surface, and it is an *(n-2)*-dimensional hypersurface where *n* is the nuclear degree of freedom of the system. Methods have been developed and are used extensively to locate the lowest point on the seam of the conical intersection, which can be regarded as a "transition state" for non-adiabatic processes.

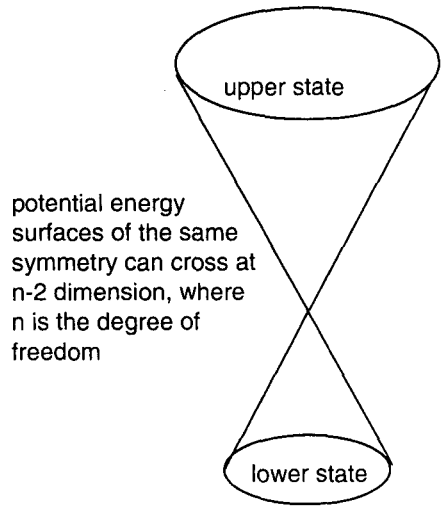


Fig. 2.3. Illustration of the concept of a conical intersection.

When heavy atoms are treated in the *ab initio* methods, the effective core potential (ECP) or pseudo-potential approximation has been proved to be very useful (Hay and Wadt 1985). In this approximation, core electrons are replaced by an effective potential. Thus, the number of electrons in the quantum mechanics is reduced,

requiring fewer basis functions and making the calculation less expensive. An additional merit of the ECP method is that the relativistic effect on valence electrons can be automatically taken into account. Thus the method is applicable to heavy atoms (second- and third-row transition metals, lanthanides and actinides) and can avoid expensive relativistic quantum mechanics calculations. Present ECPs do not permit open-shell species to be studied. Overall, ECP works very well, with very little loss of reliability and large savings. Actually the recent success of quantum mechanical calculations for catalysis, an important use of molecular modeling in industry, owes very significantly to the availability of ECP.

In summary, *ab initio* methods are the ultimate theoretical methods for electronic structure calculations, applicable to any atoms and molecules in both ground and excited states. The approximation can be systematically improved using better basis sets and better wave functions. The results can be very accurate, with an average error of ±2 kcal/mol, close to an expected experimental error. The disadvantage of *ab initio* methods is their cost, much more computationally demanding than DFT, semiempirical MO and MM methods. Especially discouraging is the high-order power dependency on the size of the system, up to N^7 or N^8. As a consequence, for example, CCSD(T) with a large basis set is routinely applicable only to molecules having at most several non-H atoms. The *ab initio* methods have been and will be used in industrial applications when the accuracy is needed or when inexpensive alternative methods, such as semi-empirical MO or DFT, do not work.

ELECTRONIC DENSITY-FUNCTIONAL THEORY (DFT)

Hohenberg and Kohn (1964) showed that the electron density, defined by:

$$\rho(r)=\rho(x,y,z)=N\int\psi^*(1,2,\cdots,N)\psi(1,2,\cdots,N)dx_2\cdots dx_N \tag{2.5}$$

uniquely determines the wave function, $\rho(r) \rightarrow V(r) \rightarrow H \rightarrow \Psi$, and the energy is a functional of $\rho(r)$:

$$E[\rho]=T[\rho]+\int\rho(r)V(r)dr+\iint\frac{\rho(r)\rho(r')}{|r-r'|}drdr'+E_{ex}[\rho]+E_{corr}[\rho] \tag{2.6}$$

To determine density, one solves the Kohn-Sham equation for "orbitals," from which the density is calculated. The fundamental problem is that the correct form of the functional is unknown. Various approximate exchange $E_{ex}[\rho]$ and correlation $E_{corr}[\rho]$ functionals have been proposed. The simplest is the local density functional approximation (LDA), in which functionals depend only on ρ. LDA was accepted widely by the physics community in the 1970s. However, this was not accurate enough for chemists for molecular modeling. Then the gradient correction (Becke 1985) was introduced in the late 1980s, in which functionals also depend on $\nabla\rho$. This made the calculated results more reliable and acceptable to chemists. Some examples of the gradient-corrected functionals are Becke88-LYP and Becke88-Perdue86. However, hybrid functionals, such as B3LYP (Becke 1993), mix the "exact (HF)" exchange with DFT in somewhat empirical fraction, improving the energetics further. These functionals have gained strong support from the chemical modeling community.

These improved DFT calculations give quite reliable energetics and properties such as NMR chemical shifts, vibrational frequencies and intensities, electrostatic potential and electron density, as good as the *ab initio* MP2 method. The Hohenberg-Kohn theorem applies only for the ground state, and there were no good DFT methods for excited states. In the last few years, many new developments are taking place for practical DFT methods for excited states, and the time-dependent perturbation method (TD-DFT) seems to be emerging as a method of choice (Jamorski, Casida and Salahub 1996).

Because of the low cost and reasonable qualitative reliability, a very large fraction of molecular modeling calculation nowadays is being performed by using DFT methods.

New developments in DFT are concerned with the search for the better functional and faster computational time. Since the exact form of the functional is not known, people have paid attention to the asymptotic form the functional should satisfy and have proposed many new functionals. Although no particular functional has so far been found to be far superior to the others, one might expect a better functional in the not too distant future. The DFT calculation is already substantially faster than the *ab initio* methods of comparable quality (such as MP2).

However, modelers want to perform calculations for even larger molecules. One approach is the development of linear scaling DFT to accomplish O(N) dependency of the computational time on the size of the system (N) (Stechel, Williams, and Feibelman 1994). Many different versions have been proposed, in both the physics and chemistry communities. Some ideas include the fast multipole method (FMM) to approximate the long-range Coulomb potential, linear scaling quadratures for exchange and correlation potential, and avoiding diagonalization of the Hamiltonian by using such methods as conjugate gradient density matrix search (CG-DMS). Nearly linear scaling for large molecules has been achieved and applied, for instance, to $(H_2O)_{1152}$ and an RNA fragment (1026 atoms, 6767 basis functions). Another approach is the "divide-and-conquer" strategy (Yang 1993), which divides a large molecule into smaller segments, performs calculations for segments, and obtains the electron density of the entire molecule by "adding" the density of segments, which is used to calculate the energy of the entire system.

Summarizing, the greatest merit of DFT methods is low cost. It could be as accurate as MP2 if proper functionals are used. They have a smaller N dependency than the comparable *ab initio* method, and they are applicable to very large molecules. Electronic properties calculated with DFT are also reasonably reliable. DFT is the method of choice for qualitative conclusions. DFT as presently implemented is often semi-empirical in nature, with a few parameters adjusted to give a best fit to some experimental results. DFT in some cases fails without any clear reason and there is no systematic way of improving the approximations. It is desirable to confirm DFT predictions by comparing them with known experimental results and high-level *ab initio* method (e.g., CCST(T)) calculations. The method for excited states is not fully established. Heavy use of DFT with gradient-corrected or hybrid functionals is being made in molecular modeling, and more industrial applications are expected in the near future.

HYBRID METHODS

Given the "pros" and "cons" of the methods described above, it makes sense to try and combine the best features of various levels and approaches. The results are the hybrid methods.

First are methods that are hybrid in a temporal sense, mixing electronic structure calculations with molecular dynamics methods (Chapter 3). For example, the so-called Car-Parrinello molecular dynamics (CPMD) method (Car and Parrinello 1985) and its relatives—e.g., DACAPO (Center for Atom-Scale Materials Physics), VASP (Kresse and Furethmüller 2000), and CAMP-Atami (Ohnishi 1994)—combine molecular dynamics (classical nuclear motion) with DFT electronic structure calculation under a periodic boundary condition with the orbitals expanded in the plane wave. This method and its various modifications have been very successful and popular for material modeling. It is being used more often now for calculations for heterogeneous and some homogenous catalysts in industrial applications.

The second class includes methods that are spatially hybrid, applying different methods in different physical regions of a molecule or computational domain. In the calculations for larger molecules and molecular assemblies (e.g., enzymes, molecules in solvated clusters, nanostructures), it is recognized that the "action" is taking place *locally* in a small, restricted region of the molecule. One should be able to combine a "more accurate (more expensive)" method for the active part and a "less accurate (less expensive)" method for the rest of the molecule. As the system to be modeled becomes larger and the accuracy required becomes higher, hybrid methods are expected to become more and more popular, in both academic and industrial applications, as the methods are tested more widely and codes become more readily available.

The so-called QM/MM method (Warshel and Karplus 1972), initiated by Warshel, Karplus and others in the early 1990s, combines quantum mechanics and molecular mechanics by dividing the Hamiltonian into the form:

$$H = H_{QM} + H_{MM} + H_{QM\text{-}MM} \tag{2.7}$$

Many other xx/MM hybrid versions have been developed. In most implementations, the combination of methods used and the applicable systems were rather specific; a specific semi-empirical MO method was combined with a specific MM method, and often additional parameters and adjustments were introduced via the $H_{QM\text{-}MM}$ term to obtain reasonable agreement with the experiment. The semi-empirical QM/MM methods have been applied mainly to biological problems such as enzyme reactions and reactions in solution. Very recently a DFT/MM method has been developed.

Such methods can be applied productively to modeling solvent effects, one of the recent hot application topics common to both *ab initio* MO and density functional methods. Most chemical measurements and experiments are performed in solution, and there is a great need to evaluate the effects of solvents on properties and reactivities of molecules. The cluster model, which includes a few explicit solvent molecules in electronic structure calculation, has been successful in describing site-specific effects, such as hydrogen bonding and donative / back donative interactions. However, for better modeling of solution, one needs a statistical treatment of solvent molecules. A hybrid QM/MM method (see below) with the statistical average of the solvent part, such as RISM-MO, has been developed. Much more is expected in this field in the near future.

Self-consistent reaction field (SCRF) models (Miertus, Scroccco, and Tomasi 1981) represent another approach, in which the solute molecule is surrounded by a polarizable continuum. Results are sensitively dependent on the shape and radius of the continuum. Thus, many different models of the reaction field and prescriptions for definition of dividing surface have been proposed. This rather inexpensive method can effectively include the bulk effect of solvent, but it misses the local site-specific effects entirely. Some efforts are being made to combine the explicit solvent model with the self-consistent reaction field model.

The chemistry of catalysis (Chapter 9) happens at "sites," physical regions where atoms or molecules bind. This is a natural situation for hybrid methods if electronic effects are most important in the immediate neighborhood of the site and when the system-wide electronic effects can be captured using a lower level of theory. Such approaches include embedded-cluster methods and electronic structure calculations with periodic boundary conditions (such as CPMD).

Another spatially hybrid technique is the ONIOM method (Svensson et al. 1996), which is based on an "extrapolation scheme" approach. In this method, the active part of the molecule is treated at the high level, while the difference between the real system and the model system, called the "substituent effect," is evaluated at the low level and the two results are combined. This method is very flexible; it can combine two MO's calculated at different levels (a unique feature) or any MO and MM (like QM/MM) in two or more layers (a unique feature), and the choice of methods and models is left to users. The integrated energy can be used to calculate energy, geometry, properties for the ground state and excited states for equilibrium structures, as well as the transition state. In general, ONIOM(CCSD(T): MP2: HF) has been recommended for high accuracy and ONIOM(DFT: HF or semi-empirical MO: MM) has been recommended for medium accuracy.

THERMOCHEMICAL CALCULATIONS

Thermochemistry includes the calculation of heat of formation, heat of atomization, proton affinity, ionization potential, and other fundamental energetics at a given temperature. Typically it is for the ideal-gas standard state. It is an important application of electronic structure theory, recently reviewed (Irikura and Frurip 1998). To make a quantitative comparison with the experiments, chemical accuracy of prediction ($<\pm 2$ kcal/mol) is required. To accomplish this accuracy, a very large basis set (such as larger than triple-zeta, multiple polarization, higher angular momentum, diffuse function) and at the same time a very high-level correlation method (such as CCSD(T), or QCISD(T)) are needed. Such a complete calculation is practical only for a molecule containing four or less non-hydrogen atoms.

Various extrapolation schemes have been proposed to achieve chemical accuracy, including G2 (Curtiss et al. 1991), G3 (Curtiss et al. 1998), CBS (Nyden and Petersson 1981), and others. In G3, for instance, at the

MP2(full)/6-31G(d) optimized geometry, the energy is extrapolated to the high-correlation, large-basis-set limit by:

$$E = E[MP4/6\text{-}31G(d)] + \Delta E(+) + \Delta E(2df,p) + \Delta E(QCI) + \Delta E(G3large) + \Delta E(SO) + E(HLC) \quad (2.8)$$

with the scaled HF/6-31G(d) vibrational frequencies for thermal correction. The so-called G2 test set consists of 148 small molecules including radicals, hydrocarbons, substituted hydrocarbons and inorganic hydrides. Using this set, the average absolute errors are 0.94 kcal/mol for the G3 method and 1.58 for G2, while DFT values include 90.88 (SVWN), 7.09 (BLYP), 20.19 (BP86), 3.11 (B3LYP) and 17.97 (B3P86) kcal/mol. The G3 theory is applicable practically to molecules with 7-10 non-hydrogen atoms. If the deviation is substantially more than average absolute error, the experiment is suspected. Reexamination of the experimental data used as a basis for such calculations is taking place at NIST and elsewhere.

Bond dissociation, breaking a bond between two atoms, is essentially a local problem, and an extrapolation scheme based on the ONIOM method is very fruitful. For instance, the H-C bond dissociation energy for H-$CR^1R^2R^3$ (9 compounds) and the C-C bond dissociation energy for H_3C-$CR^1R^2R^3$ (9 compounds), R^1, R^2, R^3 = CH_3, C_6H_5, can be calculated with the root mean square error of 2.4 kcal/mol at the ONIOM(G2MS:RMP2/6-31G*) level using H-CH_3 and H_3C-CH_3 as the model compounds, which is only slightly less accurate than the G2MS method (1.9 kcal/mol for smaller samples) and costs less than 10% of G2MS. The C-C π-bond dissociation energy of C_{60} can be evaluated with a three-layer ONIOM within 1 kcal/mol of the experiment (Vreven and Morokuma 1999; Froes and Morokuma 1999). ONIOM or similar methods will be used for calculating bond dissociation energy, the energy of reaction, and the activation energy of large molecular systems.

KINETICS CALCULATIONS

Reaction mechanisms and kinetics are also important topics for application of the electronic structure calculations. Well established are methods for determination of the transition state (TS) structure, methods for characterization of the nature of the TS by analytical energy-second-derivative calculation, and methods for following the intrinsic reaction coordinate (IRC, the minimum energy path from the TS to the reactant and the product). MP2 or hybrid DFT/DZP methods are considered to be quite reasonable for TS geometry optimization. Concerning the barrier of reaction (i.e., the energy difference between TS and the reactant or reactant complex), it is usually not difficult to determine it qualitatively. However, very accurate determination (< 1 kcal/mol) of the barrier remains challenging. One reason is that no systematic calibration against experiments has been possible due to the lack of experimental data. For another, it is expected to be harder to obtain converged results at the TS, where bond formation/breaking is taking place, than at the equilibrium structure. Reliability trends similar to those of energy of reaction (thermochemistry) calculations discussed above are usually presumed to apply, requiring a high correlation method with a large basis set.

For the calculation of the reaction rate constant using transition state theory (TST) in some gas-phase elementary reactions such as combustion reactions, reasonable theoretical estimates have been found from the energies from a G2-type scheme or from CCSD(T) with a large basis set, combined with the entropy contribution evaluated at the MP2 or B3LYP level. More sophisticated statistical theories include variational TST (Garrett and Truhlar 1979), where the TS is determined to the point of highest free energy along the one-dimensional IRC. Variational TST has been used successfully in some cases in conjunction with *ab initio* and DFT calculations to improve the agreement of the rate constant and its temperature dependency with the experiment.

For the theoretical elucidation of reaction mechanisms (i.e., the most favorable pathways on the potential energy surface of organic and organometallic reactions), qualitative calculations such as MP2 and B3LYP have been found to provide reasonable descriptions and are being applied to more and more complicated reaction systems. Semi-empirical MO methods have been used for studies of biological reactions including enzymatic reactions; however, energetics obtained at this level are marginal at best, and better theoretical methods such as DFT are being adopted more often.

SUMMARY

To summarize this chapter and give perspective, one can say that quantum mechanical methods are being used with better accuracy for modeling more complicated systems, such as "real" catalysts, chemistry in solution, large molecules (like nanostructures) and biological systems. Also:

1. Various levels of methods are likely to be used for different goals of molecular and material modeling, including semi-empirical methods (for quick semi-qualitative answers), density functional methods (for qualitative answers), and high-level *ab initio* methods (for quantitative answers).
2. Hybrid methods are likely to play a more important role in the near future. One cannot afford poor calculations for important parts (bond breaking/forming) or expensive calculations for less important parts (neighbors).
3. Combinations of quantum mechanics (electronic structure) and motion/statistics of nuclear position (MC/MD) are becoming more important.

REFERENCES

Becke, A. 1985. *Int. J. Quant. Chem.* 27. 585.

Becke, A. 1993. *J. Chem. Phys.* 98. 5648.

Car, R., and M. Parrinello. 1985. *Phys. Rev. Lett.* 55. 2471.

Center for Atom-Scale Materials Physics, *Manual for DACAPO.* http://www.fysik.dtu.dk/CAMP/dacapo.html (2000), Technical University of Denmark and University of Aarhus.

Curtiss, L.A., K. Raghavachari, G.W. Trucks, and J.A. Pople. 1991. *J. Chem. Phys.* 94. 7221.

Curtiss, L.A., K. Raghavachari, P.C. Redfern, V. Rassolov, and J.A. Pople. 1998. *J. Chem. Phys.* 109. 7764.

Dewar M.J.S., E.G. Zoebisch, and E.F. Healy. 1985. *J. Amer. Chem. Soc.* 107. 3902-3909

Dixon, S.L., and K.M. Merz, Jr. 1996. *J. Chem. Phys.* 104. 6643.

Field, M.J., P.A. Bash, and M. Karplus. 1990. *J. Comp. Chem.* 11. 700.

Frauenheim, Th., G. Seifert, M. Estner, Z. Hajnal, G. Jungnickel, D. Porezag, S. Suhai, and R. Scholtz. 2000. *Phys. Stat. Sol. (b).* 217. 41.

Froese, R.D.J., and K. Morokuma. 1999. *Chem. Phys. Lett.* 305. 419.

Fulscher, M.P., K. Andersson, and B.O. Roos. 1992. *J. Phys. Chem.* 96. 9204.

Garrett, B.C., and D.G. Truhlar. 1979. *J. Chem. Phys.* 101. 5207.

Hay, P.J., and W.R. Wadt. 1985. *J. Chem. Phys.* 82. 299.

Hohenberg, P., and W. Kohn. 1964. *Phys. Rev.* 136. B864.

Irikura, K.K., and D.J. Frurip (eds.). 1998. Computational thermochemistry: prediction and estimation of molecular thermodynamics. *ACS Symp. Ser. 677.* ACS. Washington DC.

Jamorski, C., M.E. Casida, and D.R. Salahub. 1996. *J. Chem. Phys.* 104. 5134.

Klessinger, M., and J. Michl. 1995. *Excited States and Photochemistry of Organic Molecules.* VCH.

Kresse, G., and J. Furethmüller. 2000. *VASP: The Guide.* http://cms.mpi.univie.ac.at/vasp/vasp/vasp.html, Technical (2000); http://cms.mpi.univie.ac.at/vasp/.

Laidig W.D., G.D. Purvis, and R.J. Bartlett. 1982. *Int. J. Quantum Chem. Quantum Chem. Symp.* 16. 561.

Mahapatra, U.S., B. Datta, and D. Mukherjee. 1997. 3. Recent advances in coupled-cluster methods. In *Recent Adv. Comput. Chem.* 155-181.

Miertus, S., E. Scroccco, and J. Tomasi. 1981. *Chem. Phys.* 55. 117.

Nakatsuji, H. 1979. *Chem. Phys. Lett.* 67. 329.

Nyden, M.R., and G.A. Petersson. 1981. *J. Chem. Phys.* 75. 1843.

Ohnishi, S. 1994. *CAMP Project: What is CAMP-Atami.* http://www.camp.or.jp/new/doc/camp-atami/ .

Rauhut, G., P. Pulay and Werner. 1998. *H.-J. J. Comput. Chem.* 19. 1241.

Stechel, E.B., A.R. Williams, and P.J. Feibelman. 1994. *Phys. Rev. B: Condens. Matter.* 49. 10088.

Stewart, J.J.P. 1989. *J. Comp. Chem.* 10. 209 and 221.

Stewart, J.J.P., and A.C. Stewart. 1999. Developing a semiempirical method. NATO ASI Ser. C. 539 (*Crystal Engineering: The Design and Application of Functional Solids*). 83. Kluwer.

Svensson, M., S. Humbel, R.D.J. Froese, T. Matsubara, S. Sieber, and K. Morokuma. 1996. *J. Phys. Chem.* 100. 19357.

Thiel, W. 1996. *Adv. Chem. Phys.* 93. 703.

Vreven T., and K. Morokuma. 1999. *J. Chem. Phys.* 111. 8799.

Warshel, A., and M. Karplus. 1972. *J. Am. Chem. Soc.* 94. 5612.

Yang, W.J. 1992. *Mol. Str. (Theochem).* 87. 461.

Yarkony, D. 1995. Modern electronic structure theory. *Advanced Series in Physical Chemistry*, Vol. 1 and 2, World Scientific.

Zerner, M.C. 1995. Neglect of differential overlap calculations of the electronic spectra of transition metal complexes. *Metal-Ligand Interactions: Structure and Reactivity.* NATO ASI Ser. C. 474.

CHAPTER 3

SCIENCE: MOLECULAR SIMULATIONS AND MESOSCALE METHODS

Peter T. Cummings

INTRODUCTION

Molecular simulation of fluids is having an increasingly significant impact on the chemical, pharmaceutical, materials and related industries. This is because molecular simulation provides a set of tools for predicting entirely computationally many useful functional properties of systems of interest to these industries. These properties include thermodynamic properties (such as equations of state, phase equilibria, and critical constants), mechanical properties (such as stress-strain relationships and elastic moduli), transport properties (such as viscosity, diffusion and thermal conductivity), and morphological information (such as location and shape of binding sites on a biomolecule and crystal structure). Furthermore, this list is by no means exhaustive, and it continues to grow as algorithmic and computer hardware advances make it possible to access additional properties.

Design of a chemical plant involves calculations at a macroscopic level—for example, it includes the solution of material and energy balances in a reactor and the application of computational fluid mechanics to model fluid flow using finite element methods. This macroscopic level of calculation is shown as the highest level in Fig. 3.1, which schematically attempts to summarize the relationship between the various levels of description from the most detailed electronic structure level at the bottom to the macroscale at the top.

Consider Fig. 3.1, beginning at the lowest level of description, the electronic structure/quantum mechanics level—the domain of computational quantum chemistry. Calculations at this level can predict, for example, molecular structure, free energies of formation and of reaction, dipole moments and other spectroscopic properties, and reaction rates in the gas phase and on surfaces. Computational quantum chemistry methods are described in Chapter 2.

Above this level, the Car-Parrinello method (Car and Parrinello 1985) is representative of *ab initio* molecular dynamics techniques, in which the atoms undergo motion described by classical dynamics in response to forces computed on the fly at each time step by one of the computational quantum chemistry methods, such as density functional theory. Several recent reviews highlight the application of the Car-Parrinello method (Blochl, Margl, and Schwarz 1996; Radeke and Carter 1997; Tuckerman and Martyna 2000). In particular, it has played an important role in deepening our fundamental understanding of water and aqueous solutions (Boero et al. 2000; Bruge, Bernasconi, and Parrinello 1999; Geissler et al. 2000; Marx, Sprik, and Parrinello 1997; Marx et al. 1999; Marx, Tuckerman, and Parrinello 2000; Mundy, Hutter, and Parrinello 2000; Parrinello 1997; Ramaniah, Bernasconi, and Parrinello 1998; Silvestrelli and Parrinello 1999; Sprik, Hutter, and Parrinello 1996; Trout and Parrinello 1998; Trout and Parrinello 1999).

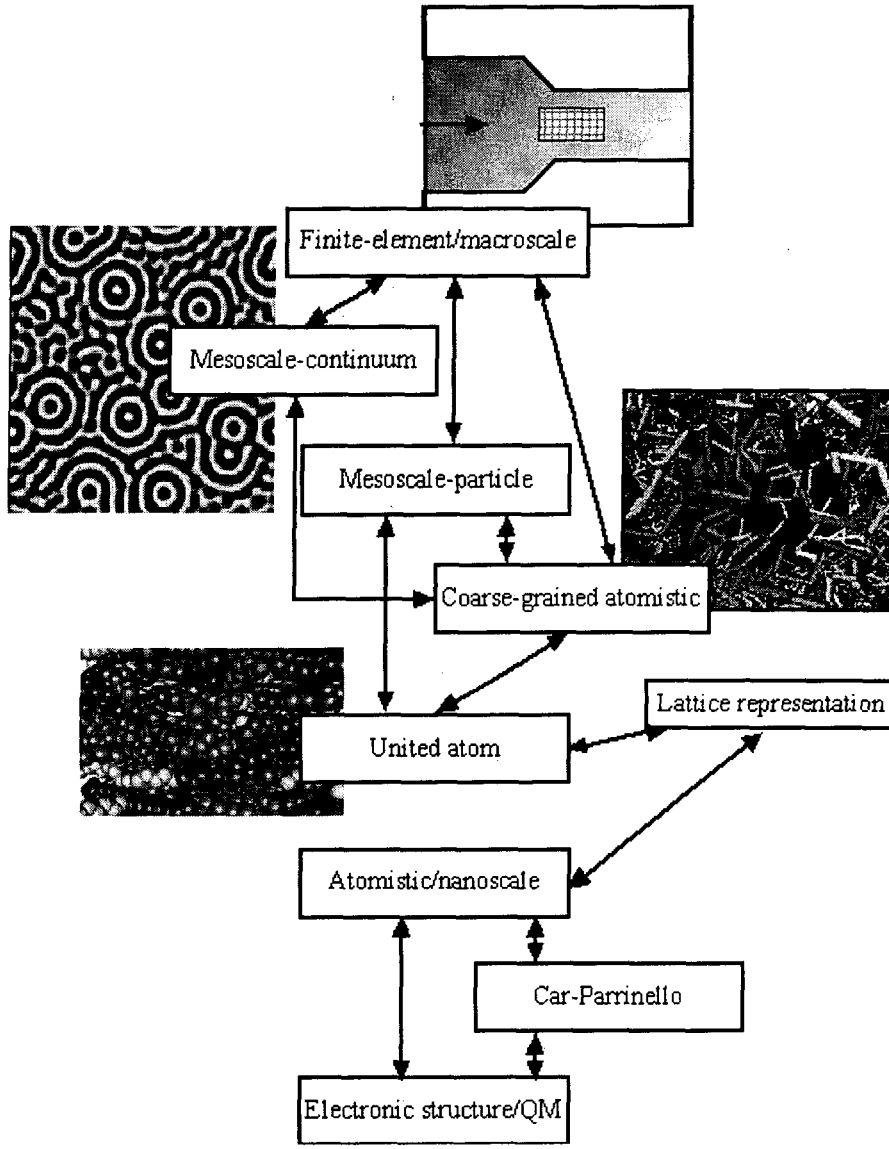


Fig. 3.1. Molecular modeling methods.

At the next highest level of description, the atomistic and united atom levels, the two main molecular simulation techniques are molecular dynamics (MD) and Monte Carlo (MC) methods, both of which have their origin and justification in classical statistical mechanics. Given mathematical models for the internal structure of each molecule (the intramolecular potential which describes the energy of each conformation of the molecule) and the interaction between molecules (the intermolecular potential which describes the energy associated with molecules being in a particular conformation relative to each other), through classical statistical mechanics, we can predict the properties of a macroscopic sample of such molecules based on statistical averaging over the possible microscopic states of the system as it evolves under the rules of classical mechanics. Thus, the building blocks are molecules, the dynamics are described by classical mechanics, and the key concept is statistical averaging. MD and MC may be summarized as follows:

- In MD, solving the classical equations of motion as a function of time (typically over a period limited to tens of nanoseconds) generates the microscopic states of the system. Thus, in MD one can observe the relaxation of a system to equilibrium (provided the time for the relaxation falls within the time accessible to MD simulation), and so determine transport properties, which at the macroscopic scale describe the relaxation of the system in response to inhomogeneities.

- In MC, equilibrium configurations are generated stochastically according to the probabilities rigorously known from statistical mechanics. Thus, MC generates equilibrium states directly (which has many advantages, including bypassing configurations which are not characteristic of equilibrium but which may be difficult to escape dynamically) and so can be used to study the equilibrium configurations of systems which may be expensive or impossible to access via MD. The drawback of MC is that it cannot yield the kind of dynamical response information that leads directly to transport properties.
- Mesoscale methods apply MD or MC techniques in domains too large for individual molecules and atoms to be included explicitly. Coarse-graining techniques are examples.

Central to both the MD and MC methods is the need for accurate force fields to describe the inter- and intra-molecular interactions, which today are determined by a combination of computational quantum chemistry methods and fitting to available experimental data. The development of these force fields is a very active sub-field of molecular simulation.

In this chapter, we review some of the major developments in molecular simulation that have made it possible to apply these techniques to the complex systems encountered in the chemical, pharmaceutical, materials and related process industries. Two recent textbooks provide excellent reviews of molecular simulation methods (Allen and Tildesley 1987; Frenkel and Smit 1996).

MOLECULAR SIMULATION METHODS

Canonical Ensemble MD or MC

The most common form of MD or MC simulation is that conducted on a system at fixed temperature (T), volume (V) and number of molecules (N). This is frequently referred to as a canonical ensemble simulation in view of the fact that this is the name given to the NVT ensemble in statistical mechanics. It is also less frequently referred to as the isothermal-isochoric ensemble. In an NVT ensemble simulation, whether MC or MD, one can predict the pressure (P), internal energy (U), structure (to make contact with neutron and X-ray scattering experiments), and properties defined as fluctuations (such as specific heat). In MD (but not MC) simulations, linear response transport properties such as diffusivity (D), thermal conductivity (λ), and viscosity (η) can also be calculated. NVT simulations have been performed on wide variety of systems, including polymers (Binder 1997; Binder and Paul 1997; Escobedo and de Pablo 1999; Freire 1999; Hofmann et al. 2000; Uhlherr and Theodorou 1998), proteins (Biggin and Sansom 1999; MacKerell et al. 1998; Neumaier 1997; Onuchic et al. 2000; Pande, Grosberg, and Tanaka 2000; Rognan 1998), and meso- and nano-structured materials (Barton et al. 1999; Cracknell et al. 1995; Gelb et al. 1999; Zachariah and Carrier 1999).

Closely related to the NVT ensemble is the isothermal-isobaric, or NPT ensemble, in which the pressure P, rather than the volume, is held fixed (Allen and Tildesley 1987; Andersen 1980).

Essentially all commercial molecular modeling codes feature one or both of these ensembles. NVT and NPT are popular ensembles for many biological calculations of protein structure. Many public domain NVT and/or NPT codes are available. The most widely used is the MD_POLY program produced by Daresbury Laboratory in the United Kingdom (see site report, Appendix B) which permits both NVT and NPT simulations of fluids composed of complex molecules (i.e., polyatomic molecules).

In measuring the complexity of NVT and NPT MD and MC calculations, one MD time step is typically of the order of 10^{-15} s. In the absence of specialized methods for improving the efficiency of MC (such as continuum configurational bias for chain molecules (de Pablo et al. 1993; Frenkel and Smit 1996)), one MD step is comparable to the computational expense of approximately N MC steps, where N is the number of molecules in the simulation, and is frequently referred to as a MC cycle. A good measure of complexity for MD is the number of MD atom-steps $N_{atom\text{-}steps}$, which is the product of the number of atoms in the simulation N and the number of time steps $N_{time\text{-}steps}$, *viz.*

$$N_{atom\text{-}steps} = N \times N_{time\text{-}steps} \tag{3.1}$$

The range for $N_{atom\text{-}steps}$ in published MD simulations today is approximately 10^6 to 10^{13}, with the high end represented by the 1-microsecond folding of a small protein in water (Duan and Kollman 1998). Modern personal computers and workstations can handle computations comfortably up to $N_{atom\text{-}steps} \cong O(10^8$ to $10^{10})$. Beyond this, parallelization becomes necessary to complete calculations within a reasonable time.

There are several strategies for parallelizing molecular dynamics calculations onto many-processor systems (Cochran et al. 1998; Plimpton 1995; Plimpton and Hendrickson 1995; Plimpton and Hendrickson 1996). Broadly speaking, there are two classes: data parallel methods and domain decomposition.

In one type of data parallel method, the replicated data method, each processor stores a copy of quantities associated with all N atoms (positions, momenta, and higher derivatives as required by the numerical integration method of choice) in the simulation. The force calculation is divided equally across processors to achieve load-balance. The results of the force computation as well as the new updated positions of the atoms must be communicated to all processors at every time step. The cost for this global exchange of information scales as the number of processors P, and so is the reason replicated data does not scale well to large numbers of processors. An enhancement to replicated data, known as force-decomposition (Plimpton 1995; Plimpton and Hendrickson 1996), reduces this communication cost by a factor of $1/\sqrt{P}$, allowing more processors to be used effectively on a given size problem. However the non-local communication in the method still begins to dominate as machine sizes grow to thousands of processors.

In domain decomposition, the spatial domain of the central simulation cell is subdivided into cubic sub-domains, each of which is mapped to a processor (Pinches, Tildesley, and Smith 1991). Thus each processor is responsible for integrating the equations of motion for the molecules lying within its spatial subdomain. Domain decomposition scales well for very large numbers of molecules (large N) and hence is the method of choice in simulations with large numbers of molecules to be simulated for short real times (1-100 ps, for example). Hard materials simulations frequently fall into this latter category (Nakano, Kalia, and Vashishta 1999).

Data parallel methods work best for relatively small systems that must be simulated for long real times (1-100 ns and beyond). An example is the simulation of lubricants over periods of 10 ns and longer (Cui et al. 1998; McCabe et al. 2001; McCabe, Cummings, and Cui 2000; Moore et al. 2000a; Moore et al. 2000b; Moore et al. 1999; Moore et al. 1997).

Thus far, parallelization has not received much emphasis in commercial simulation codes, and in such codes, it is limited to data parallel forms. This is in part because few industrial users of molecular simulation codes have access to the large massively parallel computers that are typically used in domain decomposition simulations. However the increasing adoption of Beowulf clusters by companies, one of the hardware trends observed as part of this study, may in the future lead to industrial demand for commercial and/or public domain-decomposed codes.

Transport Properties and Nonequilibrium Molecular Dynamics

Transport properties, such as viscosity, diffusivity and thermal conductivity, can be calculated from MD simulations. This should not be surprising because transport coefficients are the quantities which describe the way in which a macroscopic system relaxes in time away from some inhomogeneity or non-equilibrium state. Although an equilibrium MD simulation is at equilibrium on average, there will be fluctuations away from equilibrium on the microscopic scale. By measuring the autocorrelation of these fluctuations, one can measure the relevant transport property (Zwanzig 1965).

A succinct review of the formulae required for implementing equilibrium MD techniques to compute transport properties in fluids of spherically symmetric molecules is given by McQuarrie (McQuarrie 1976). There are basically two kinds of formulae: Green-Kubo (GK) formulae, in which the transport property is expressed in terms of an integral of the relevant autocorrelation function, and the Einstein formulae, which in essence are integrated forms of the GK relations. Einstein relations can be derived directly by considering microscopic expressions for quantities (such as the mass density and momentum density), substituting these expressions into the corresponding macroscopic transport equation (such as Fick's law of diffusion or the Navier-Stokes equation incorporating Newton's law of viscosity), and solving for the transport coefficient of

interest (such as diffusion coefficient and shear viscosity). As an example of the GK and Einstein expressions, consider diffusivity: The Einstein relation for the self-diffusion coefficient D_s is given by

$$D_S = \lim_{t\to\infty} \frac{\left\langle \sum_{i=1}^{N} [r_i(t) - r_i(0)]^2 \right\rangle}{6t} \tag{3.2}$$

where $\langle ... \rangle$ indicates equilibrium ensemble average and $\mathbf{r}_i(t)$ is the position of the center of mass of molecule i. The GK relation for D_s is:

$$D_S = \frac{1}{3m^2} \int_0^\infty \left\langle \frac{1}{N} \sum_{i=1}^{N} \mathbf{p}_i(t) \cdot \mathbf{p}_i(0) \right\rangle dt \tag{3.3}$$

where $\boldsymbol{p}_i(t)$ is the momentum of molecule i at time t so that $\left\langle \frac{1}{N} \sum_{i=1}^{N} \mathbf{p}_i(t) \cdot \mathbf{p}_i(0) \right\rangle$ is the momentum autocorrelation function. Similar Einstein and GK equations exist for viscosity and thermal conductivity. Note that the Einstein relation involves quantities $\boldsymbol{r}_i(t)$, which are obtained by integrating the quantity in the autocorrelation function [$\boldsymbol{p}_i(t)$]. This is a common feature of Einstein and GK relations, and in fact Eq. 3.3 can be obtained by differentiating Eq. 3.2 or vice versa.

While equilibrium methods are used extensively for calculating transport properties (Cummings and Evans 1992), in recent years non-equilibrium molecular dynamics (NEMD) methods have become popular for studying transport processes and, in particular, those related to viscosity (the rheological properties). This is because equilibrium MD methods can only access the linear regime of transport properties, while NEMD can be used to study transport properties in the non-linear (i.e., non-Newtonian in the case of rheology) regime. The general philosophy of the NEMD method (Cummings and Evans 1992; Evans and Morriss 1990; Sarman, Evans, and Cummings 1998) is to introduce a (usually fictitious) field X into the equations of motion of the system, which drives the conjugate thermodynamic flux J (such as the momentum flux in planar Couette flow or the heat current for thermal conductivity). The first requirement for this applied field is that it be consistent with periodic boundary conditions to ensure that the simulation sample remains homogeneous. The second requirement is that the transport property δ of interest can be calculated from the constitutive relation:

$$\delta = \lim_{X\to 0} \lim_{t\to\infty} \frac{J}{X} \tag{3.4}$$

The proof that an algorithm satisfies these two requirements is given by linear response theory (Evans and Morriss 1990). In many cases the value of the transport coefficient at non-zero fields has no physical meaning because the field used to drive the thermodynamic flux is fictitious. However, the NEMD algorithm for planar Couette flow (Evans et al. 1983; Ladd 1984) used to calculate shear viscosity can be shown to be exact for all values of the field (which in this case is the velocity gradient represented by the strain rate $\gamma = \partial u_x / \partial y$, the rate of change of the streaming velocity in the x-direction with vertical position). While NEMD methods have been in use for nearly three decades, recent theoretical and algorithmic developments have placed the methods on a firm foundation (Evans and Morriss 1990; Sarman, Evans, and Cummings 1998).

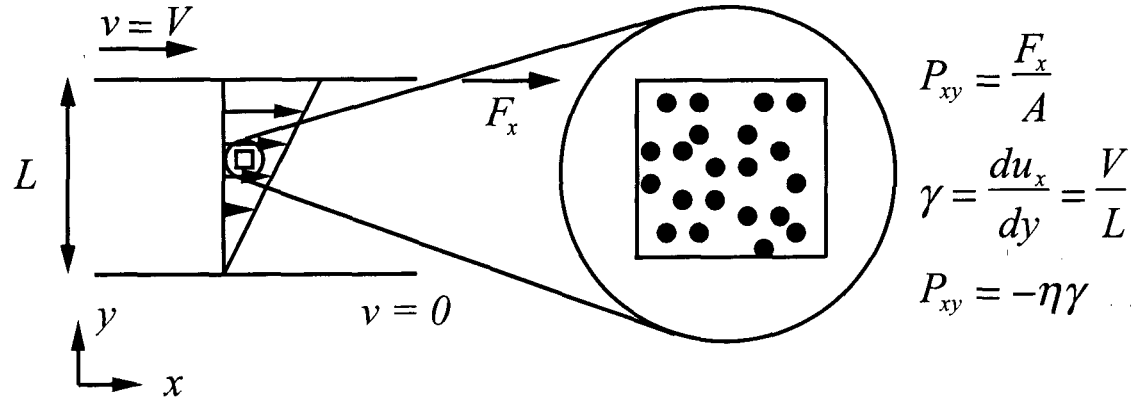


Fig. 3.2. Schematic of NEMD simulations of planar Couette flow. Lees-Edwards boundary conditions (Lees and Edwards 1972) are used to ensure that the simulation sample is sheared consistent with a macroscopic strain field, as shown here.

Recent examples of the use of NEMD include calculating the viscosity index of several lubricant components (McCabe, Cummings, and Cui 2000; Moore et al. 2000c; Moore et al. 1997) and the steady-state and dynamic rheological properties of short polyethylene melt (Moore et al. 2000a; Moore et al. 2000b; Moore et al. 1999).

Grand Canonical Monte Carlo (GCMC)

In the grand canonical Monte Carlo method (Allen and Tildesley 1987; Norman and Filinov 1969), simulations are performed in the grand canonical ensemble of statistical mechanics, in which the chemical potential of the molecules (μ), temperature (T), and volume (V) are fixed. For mixtures, the chemical potentials of each species are also fixed. GCMC has many useful applications because it models a system that is in equilibrium with a bulk fluid at a specified chemical potential. Molecule insertions and deletions are used to achieve equilibrium with the bulk fluid. Examples of applications where it would be desirable to have a system in equilibrium with a bulk fluid include the following:

- Solvent molecules between two large colloidal particles in equilibrium with the bulk solvent
- Fluid adsorbed on a surface or in a pore in equilibrium with a bulk fluid phase
- A fluid phase in equilibrium with another fluid phase

Thus, GCMC can be used in calculating vapor-liquid phase equilibria if the chemical potential of one of the phases is known. Alternatively, it is possible to iterate on two GCMC simulations (one of liquid, the other of vapor) in which the chemical potential is varied by trial and error until the pressure of the two GCMC simulations is the same.

GCMC has the negative feature that it fails or becomes difficult to converge at high density where the probability of an insertion being successful is low. Nevertheless, it is a very useful technique and has been applied extensively. For a review of its application in adsorption, see Gelb et al. (1999).

Gibbs Ensemble Monte Carlo

One of the very significant recent developments in molecular simulation was the development of the Gibbs ensemble Monte Carlo (GEMC) method by Panagiotopoulos in 1987 (Panagiotopoulos 1987a; Panagiotopoulos 1987b; Panagiotopoulos et al. 1988). The GEMC method allowed for the first time the relatively routine calculation of phase equilibria for both pure fluids and mixtures. Furthermore, the critical properties predicted by a given force field could be calculated, and the force field tuned to reproduce the phase envelope and critical point accurately. The technique, illustrated in Figure 3.3, involves the direct simulation of two fluid phases in equilibrium, though not in physical contact. It is to be contrasted with a brute-force method in which the system is quenched into the two-phase region with a subsequent separation

into two phases. Such brute-force methods, also known as quench MD (since usually MD is the underlying simulation method used), are useful if one of the foci of the study is the interface between the two coexisting phases (e.g., to calculate the surface tension between the two phases). If the sole goal is to know the densities, compositions and pressure of the two coexisting phases at equilibrium, then GEMC is more efficient because no interface is present in the simulation. Instead, the two boxes in the simulation represent samples from the bulk regions of the coexisting phases.

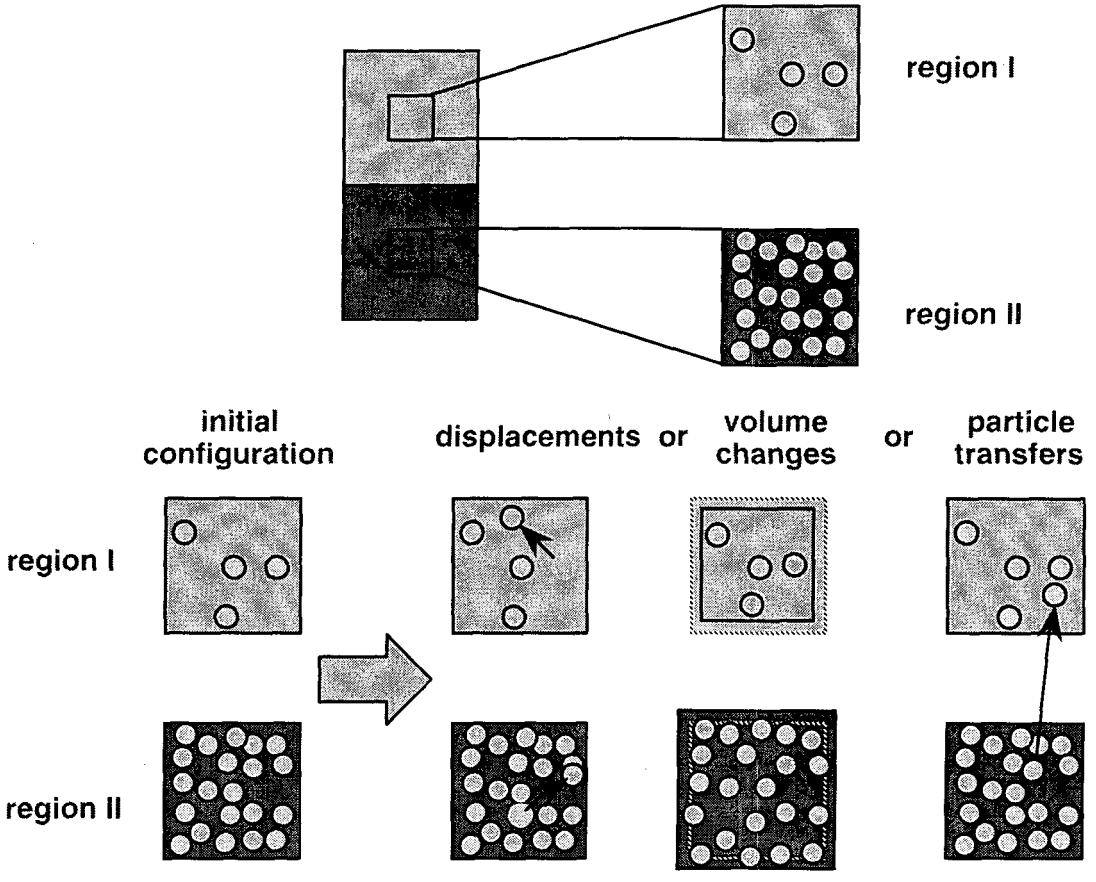


Fig. 3.3. Schematic of GEMC. Two simulation cells, each corresponding to a fluid phase, are brought to phase equilibria by a combination of displacement, volume exchange and molecule exchange moves with acceptance criterion designed to ensure that the conditions of phase equilibrium are met (after Panagiotopoulos 1987b).

Hence, the primary advantage of GEMC is that its use of two boxes with molecule exchange directly yields the correct equilibrium point for two phases in coexistence. Another advantage is that it is an efficient method for mixtures as well as pure fluids, and in fact for mixtures, the computations can be performed at constant pressure. The main drawback is that, just as with the insertion method for GCMC, GEMC fails at high densities. Furthermore, near the critical point, there are large fluctuations in the density which can result in the two boxes switching identity several times in the course of the simulation (i.e., the lower and higher density boxes switch to being the higher and lower density boxes respectively). GEMC has been applied to bulk fluid phase equilibria, phase equilibria of fluids confined to pores, and can be combined with the Gibbs-Duhem method (Kofke 1993; Mehta and Kofke 1993) to map out full phase diagrams.

GEMC and other methods for predicting phase equilibrium are reviewed in several recent review articles by Panagiotopoulos and co-workers (Orkoulas and Panagiotopoulos 1999; Panagiotopoulos 1992; Panagiotopoulos 1996; Panagiotopoulos 2000).

Kinetic Monte Carlo

Kinetic Monte Carlo models reaction kinetics within a molecular simulation framework by stochastic transitions (using rate constants) (Gillespie 1968; Gillespie 1976). It allows for the direct incorporation of *ab initio* rate constant results and permits the incorporation of non-local phenomena. Several reviews of application areas in which kinetic Monte Carlo has had a significant impact, including rational catalyst design, have recently appeared (Levi and Kotrla 1997; Martin 1998; Stoltze 2000).

Histogram Reweighting

Histogram reweighting proposed by Ferrenberg and Swendsen (Ferrenberg and Swendsen 1988; Ferrenberg and Swendsen 1989) involves computing a density distribution in the grand canonical μVT ensemble and inferring from this the densities of the coexisting phases. Moreover, by taking advantage of the way in which the temperature enters into the expression for the probability distribution for the grand canonical ensemble, one can obtain the density distribution at nearby temperatures and chemical potentials from the same simulation run. Hence, it provides an efficient way to trace out a complete phase diagram. Though initially developed by Ferrenberg and Swendsen for lattice models of magnetic systems, the histogram reweighting method has been generalized to continuum models of fluids (Orkoulas and Panagiotopoulos 1999; Panagiotopoulos 2000). Several reviews of this technique are available (Binder et al. 1995; Swendsen, Wang, and Ferrenberg 1995).

Continuum Configurational Bias Monte Carlo

Both in GCMC and GEMC, it can become essentially impossible to insert randomly a long-chain molecule (e.g., an alkane) into an even moderately dense liquid phase. Additionally, even in the case of simply equilibrating an NVT simulation of chain molecules, one faces the problem of efficiently moving the molecules around in order to reach equilibrium at a reasonable pace. The continuum configurational bias method (CCBMC) method is a technique which can be used to address both of these difficulties (de Pablo, Laso, and Suter 1992; de Pablo et al. 1993; Laso, de Pablo, and Suter 1992; Mooij, Frenkel, and Smit 1992). For insertions, the CCBMC method involves placing the molecule into the liquid phase one monomer at a time, choosing the location of each new monomer in an optimal fashion. This process introduces a bias into the insertion procedure, which must be compensated for in the acceptance criteria (Frenkel and Smit 1996). Likewise, to move a molecule within an MC NVT simulation, CCBMC methods include deleting and reinserting all or part of a molecule, restoring the molecule into the liquid phase one monomer at a time, choosing the location of each new monomer in an optimal fashion, and then correcting for the introduced bias in the acceptance criterion. The CCBMC method has made it possible to perform efficient MC and GEMC simulations of polymer systems.

Multiple-Time-Step MD

In molecular dynamics simulations, one is essentially solving a large number of simultaneous nonlinear differential equations. For a fluid of monatomic molecules, there is a single characteristic time in the system, the translational relaxation time, and so the choice of time step is straightforward. For complex molecular systems, such as alkane chain molecules, there are several characteristic times: the translational relaxation time, the rotational relaxation time, and the times characterizing C-C bond-stretch, C-C-C angle-stretch and C-C-C-C torsional motion (i.e., the intramolecular motion). The intramolecular motion is typically on a time scale that is one order of magnitude smaller than the intermolecular motion. This leads to stiff systems of nonlinear differential equations, which if treated naively would require the use of a time step small enough to capture the fastest motion, leading to inefficient simulations. Tuckerman et al. (Tuckerman, Berne, and Martyna 1992) developed a theoretically-based reversible multi-time-step algorithm that uses two or more time steps, each appropriate to one mode of motion (e.g., translation and bond bending/stretching). This method has been used extensively by many research groups in performing simulations of chain molecules. It was extended to NEMD simulations by Cui and colleagues (Cui, Cummings, and Cochran 1996).

Reactive Monte Carlo

Reactive Monte Carlo is a technique that permits the incorporation of chemical equilibrium into MC simulations. It can be considered to be an application of GEMC concepts in that it uses insertion/deletion/change in identity moves to impose the condition of chemical equilibrium on the reactant and product species (Johnson, Panagiotopoulos, and Gubbins 1994; Smith and Triska 1994). The method (1) can be combined with GEMC to evaluate the impact of phase equilibria on chemical equilibria and (2) with GCMC modeling of adsorption to probe the impact of confinement to pores on chemical equilibrium.

ISSUES

MD, MC and their variants are routinely applied, both industrially and in academic research, to a wide variety of systems, ranging from low-molecular-weight inorganic and organic fluids to biological systems and polymers, and with a high degree of success. Nevertheless, there are several outstanding issues in the field of molecular simulation that have been identified as a result of this study. Many of the issues have been raised in prior studies and workshops on molecular simulation.

Force Fields

Force fields are relatively well developed for biological conditions where the focus is on a narrow range of state conditions (ambient and near-ambient). By comparison, they are poorly developed for chemical process conditions, where the demands of predicting phase equilibria require accuracy over wide ranges of density and temperature. To develop a "chemically useful" force field (one that can be used to predict phase equilibria in very good agreement with experimental data) requires fitting the potential parameters to the phase envelope, not just the physical properties at ambient conditions. Today we have chemically useful force fields for a limited range of systems, including alkanes, alkenes, perfluoroalkanes, CO_2, H_2O and other low molecular weight species. In order to develop a large enough repertoire of chemically useful force fields for the almost-infinite variety of systems encountered in the chemical industry, better and more rapid methods for force field development are required. In particular, rapid prototyping methods are needed to predict reasonably accurate force fields for new materials.

At this time, several research groups are independently developing force fields for various classes of fluids—e.g., the alkanes (Chen and Siepmann 1999; Errington and Panagiotopoulos 1998; Martin and Siepmann 1998; Nath, Escobedo, and de Pablo 1998; Wick, Martin, and Siepmann 2000) and water (Chialvo and Cummings 1999). Coordinated research efforts aimed at dividing the effort to develop chemically accurate force fields are desirable, but they do not exist today. A related problem is that the development of force fields is not regarded as a "glamorous" research activity, and government funding is difficult to obtain solely to develop force fields. Typically, force-field development is undertaken as part of broader projects whose aims are to simulate specific systems accurately. While commercial vendors have undertaken significant force-field development activities, their efforts have not focused on the accurate prediction of phase equilibria as a design goal for the force field. There has been substantial effort on force-field development for applications in the pharmaceutical industry, where the range of conditions to be simulated is far narrower than in chemicals/materials applications. Finally, many users are not aware of the limitations of existing force fields used in commercial molecular modeling packages because the commercial packages do not currently permit calculation of phase equilibria. In fact, many calculations in commercial packages are performed in the *NPT* ensemble, in which the output is the energy and the density, both quantities being relatively insensitive to force field errors.

Time Scales and Mesoscale Modeling

An enduring difficulty in MD is the time scale accessible to MD simulations. Dill (Chan and Dill 1993) summarized the nature of the problem in a 1993 *Physics Today* article. Straightforward classical molecular dynamics simulation will not today, nor for the next decade or two, span the time scales present in complex molecular systems (from 10^{-15} s to 10^{1} s and beyond) even with a completely efficient implementation on the world's fastest parallel supercomputers. The essential problem of simulating over long periods of time relative to the shortest time-scale motion is not fundamentally changed, even by recent extraordinary

advances in parallel supercomputing. While parallel supercomputing makes it possible to span larger spatial scales rather easily, spanning larger time scales remains a fundamental difficulty.

The time-scale problem is partially solved by MC, which enables transitions between configurations that would be possible in MD only by extraordinarily long time trajectories. For example, for one polymer molecule to pass through another molecule on the route to an equilibrated configuration would take many MD steps but can be achieved in a single CCBMC step. As noted above, an order of magnitude increase in the time accessible to MD simulations can be achieved through the use of multiple-time-step methods, but this is a relatively small gain. For systems in which there are bottlenecks isolated in time, so-called rare event simulations, algorithms have been developed to address the efficient simulation of such systems (Sorensen and Voter 2000; Voter 1997a; Voter 1997b; Voter 1997c; Voter 1998). However, such algorithms do not address systems that simply require long execution times to simulate beyond the relaxation time of the system, such as polymer melts of very long chains or protein folding.

One approach that has seen rather limited application is the development of special-purpose MD hardware, parallel supercomputers composed of processors whose microcode directly supports the MD code. This approach has been more popular in Japan than in the U.S. or Europe, as is evidenced by the site reports included in Appendix B of this report (see JRCAT, NEC, RIKEN reports).

Finally, it should be pointed out that one of the most popular approaches to extending MD time scales is to simulate a coarse-grained version of the molecule of interest. For example, as described in the section entitled, "Polymers and Mesoscale Modeling" below, a common coarse-grained model for a polymer is the bead-spring model, which can be simulated on much longer time scales than a fully atomistic or united-atom model of the same polymer. In the case of a polymer solution, Brownian dynamics (Ermak and McCammon 1978) or some variation such as Stokesian dynamics (Phung, Brady, and Bossis 1996) is used to account for the drag and hydrodynamic effects of the solvent. Although bead-spring models extend the time scales over which polymer dynamics can be modeled, the relationship between a bead-spring model and the real polymer being modeled is somewhat obscure.

In order for coarse-grained simulations to be able to meaningfully reflect differences in chemical composition and architecture of the molecules and systems being approximated, rigorous methods are needed for coarse-graining a detailed model or restoring the detail from a coarse-grained model to its atomistic counterpart. One such mapping is provided by Mattice and co-workers (Haliloglu et al. 1996; Haliloglu and Mattice 1998), who provide a unique mapping between an explicit atom model for a polymer in a continuum space and a united atom model on a highly connected lattice. The problem of addressing the multiple time scales in polymers and rigorously moving between the varying levels of detail in describing the molecules, all within a single integrated simulation package, is the subject of the Doi Project in Japan, profiled in the site reports (see Nagoya University—"The Doi Project").

Dissipative particle dynamics (DPD), in which particles represent mesoscopic elements of the underlying fluid, is emerging as one of the most popular coarse-grained methods in both academia and industry (Bonet, Avalos, and Mackie 1999; Evans 1999; Flekkøy and Coveney 1999; Groot, Madden, and Tildesley 1999; Groot and Warren 1997; Marsh and Coveney 1998; Schlijper, Hoogerbrugge, and Manke 1995). The dissipative particles evolve under a dynamics similar to molecular dynamics but with different inter-particle forces. The connection between DPD and the macroscopic (hydrodynamic, Navier-Stokes) level of description has been made (Espanol 1995; Groot and Warren 1997), and very recently, the connection between DPD and atomistic/nanoscale molecular dynamics has been developed (Marsh and Coveney 1998). Note that much of this development has been by Groot, Warren, and Tildesley, industrial researchers at Unilever.

Error Estimation

As molecular simulation emerges from a purely academic discipline to one with an increasing industrial relevance, one of the key questions to emerge will be: How accurate is a given calculation? Routinely, molecular simulation results report the statistical error associated with the calculation itself. However, the intrinsic error of the force field model (and method) is rarely, if ever, reported.

As a specific example, a user of a particular physical property (e.g., the critical temperature and pressure of an as yet unsynthesized compound) always needs to know the error associated with that property in order to propagate its effect on the prediction of quantities dependent on that property. For such a user, it is then critical to know what is the error associated with the critical temperature and pressure derived from the molecular simulation. Today, such estimates can be obtained for homologous series of compounds by comparing with predicted values against available experimental data for some members of the series. In a few cases, calibration of the force field itself is possible by comparison to electronic structure calculations (e.g., water), and the effect of the force field error on a computed physical property can then be estimated by sensitivity analysis.

While error estimation is rare today, in the future the need for intrinsic error estimation will presumably lead to the development of better methods to estimate error.

Molecular Simulation Technology Transfer

Significant differences between the way in which the molecular simulation community and the computational quantum chemistry fields are structured emerged early in this WTEC panel's workshop discussions. Compared to the fairly narrow focus of a quantum chemistry computation, the objective of molecular simulation can be very broad. For example, the objective can vary widely from the calculation of simple thermodynamic properties (pressure and energy of a bulk fluid, for example) to the phase envelope of a fluid confined to a porous medium, from bulk transport properties (such as diffusion and thermal conductivity) to the non-Newtonian nanorheology of lubricants confined to nanoscale gaps in micromachine and hard disk drive applications.

The result is that a very large number of research groups around the world, based in a broader range of disciplines (including physics, chemistry, chemical engineering, biology, biochemistry, and geochemistry), constructively contribute to the refinement of existing molecular simulation methods (such as Monte Carlo, molecular dynamics, and their intermediate hybrids) or development of new molecular simulation techniques. A small but increasingly important component of the molecular simulation research community is developing intermolecular potentials for use in molecular simulation applications. The latter activity is especially important as molecular simulation techniques are maturing to the point of predicting properties of practical importance to chemical process and product design activities (such as critical point properties, vapor pressure curves, and lubricant viscosity).

However, within both the molecular simulation methodology and intermolecular potential development communities, there is a dichotomy between those whose focus is primarily on biological systems (where the primary concern is the accurate prediction of properties in a rather narrow range of temperature and pressure near ambient conditions) and those whose focus is on systems of broad interest in the chemical industry (and so are concerned with accurately predicting properties over wide ranges of temperature and pressure including phase equilibria). One important consequence of the relatively large and diverse pool of developers of molecular simulation methods and intermolecular potentials is that adoption of new molecular simulation methodologies and intermolecular potentials by industry and third-party software vendors is significantly slower than for the corresponding computational quantum chemistry community.

The fractured nature of the molecular simulation field has resulted in third-party vendors having an understandable reluctance to incorporate the latest molecular simulation developments and intermolecular potentials into their commercial packages. As a result, the methods available to industry on a routine basis lag significantly behind those available to academic and national laboratory researchers—behind by as much as a decade or more. Clearly, there is need for new paradigms for timely distribution of the latest molecular simulation capabilities in a form useful to industry, academia and national laboratories.

Polymers and Mesoscale Modeling: A Particular Challenge

Polymers represent a class of physical materials of obvious interest to the chemicals and materials industry. Polymeric materials are among the products of the chemical industry with the largest volume, largest revenue earnings, and most success. From a molecular modeling point of view, polymers represent a particular challenge. First, small differences in polymer architecture or chemical composition can give rise to very

different properties. Thus, some properties can have a strong dependence on the detailed molecular structure. For example, most polymer mixtures (or blends) are immiscible. This is because small differences between the repeat groups in the polymer become magnified into very large differences between polymer molecules when the molecules contain thousands or more repeat units. Thus, while low-molecular-weight versions (a few monomeric units) of two polymers might be miscible, a blend containing large-molecular-weight versions of the same molecule is likely to be immiscible. Molecular modeling methods need to be able to reflect this sensitivity to molecular architecture and chemistry.

On the other hand, for a given architecture and chemistry, a polymer exhibits scaling properties with size that must also be captured by molecular modeling methods. The existence of such extremes, extreme sensitivity in some properties, absence of dependence on anything but a few variables such as size in others, makes the molecular modeling of polymers extremely challenging, and has led to the development of methods applicable at various levels of description, ranging from fully detailed atomistic models to models in which all but the most essential elements have been coarse-grained away (as in dissipative particle dynamics models of polymers). Some of these techniques have been described above in the section entitled, "Time Scales and Mesoscale Modeling."

The essential difficulty in developing molecular modeling methods for polymers is their long relaxation times, which can range from nanoseconds to microseconds, seconds, hours and even days or months, depending on molecular weight. For a linear polymer, the relaxation times grow with n^2, where n is the number of monomeric units in the polymer. For all but the shortest polymers, this renders the relaxation times well beyond the range accessible to molecular dynamics. For example, for short polyethylene ($C_{100}H_{202}$), the rotational relaxation time in the melt at 448 K is about 3.5 ns (Moore et al. 2000a), which means that a molecular dynamics simulation of $C_{100}H_{202}$ must be run for several rotational relaxation times to equilibrate, and then sampled over several relaxation times to ensure that the averages measured represent averages over a representative sample of equilibrium configurations. Such calculations have a complexity $N_{atom\text{-}steps} \approx 10^{11}$ atom-steps, and so require significant levels of parallel resources. As noted earlier in this chapter, several strategies have been developed to work around the relaxation time problem. We summarize those approaches here in the context of their polymers applications. If the computation is aimed at predicting equilibrium non-dynamical properties (i.e., properties that do not require dynamical information such as transport properties), MC can be a much more efficient route to equilibrated configurations over which equilibrium properties can be averaged. The CCBMC methods have extended dramatically the range of polymer systems accessible to simulation by Monte Carlo methods. Added to the CCBMC methods are a variety of pseudo-ensemble and tempering techniques that further extend the range of systems accessible to molecular simulation. These developments are the subject of recent reviews (de Pablo and Escobedo 1999; de Pablo, Yan, and Escobedo 1999).

Whether the properties to be calculated are dynamical or not, coarse-graining the polymer at one of several levels of description is a popular strategy. From the fully atomistic model (explicit-atom model, EA, or explicit-hydrogen model, EH), the first level of coarse-graining is to use a united-atom (UA) model, in which the hydrogen atoms are not treated explicitly. Instead, spheres represent the CH_3, CH_2, and CH functional groups and the C-C bond-stretch, C-C-C angle-stretch and C-C-C-C torsional motion are all retained. This reduces computation times by one or more orders of magnitude depending on the size of the molecules and the method being used (MD or MC). A variation on this approach is that of Mattice and co-workers (Haliloglu et al. 1996; Haliloglu and Mattice 1998), who take an EA model in continuum space, map it uniquely onto a UA model on a highly connected lattice, equilibrate using lattice methods, then use the unique inverse mapping to produce an equilibrated EA configuration.

The next level of coarse-graining is typified by bead-spring models, in which clusters of monomer are replaced by a bead (or center of force) connected by springs. Because the fastest intramolecular motions (bond and angle stretch) have been eliminated, a bead-spring model for a polymer melt or solution can be simulated on much longer time scales than a fully atomistic or united atom model of the same polymer. The problem with bead-spring models is the lack of a rigorous mapping between the beads and the spring forces and the corresponding UA or EA model. Bead-spring models were first developed to describe a single polymer molecule in a Newtonian solvent, in which case the mapping between the polymer and the bead-spring model is well understood (Bird et al. 1987). In this case, it is clear that a bead represents 20 or more monomeric units of a polymer, and the diameter of the bead the effective hydrodynamic radius of this group

of monomers. The consequence is that the maximum extension of the spring can be many times the diameter of the bead. When bead-spring models were adopted to represent polymer melts (Kremer and Grest 1990; Rudisill and Cummings 1991), it was realized that such large bond lengths would lead to bead-spring polymers being able to pass through each other physically, and so the bond lengths were made short enough to prohibit this possibility. In this process, the connection to the underlying UA or EA model is evidently lost. (This is not true in the Mattice et al. mapping, because the lattice imposes rotational isomeric state configurations on the model, thus preserving the connection with the underlying EA model.) As a consequence, bead-spring models are mostly used to examine trends, typically the impact of increasing length or degree of branching of the polymer molecule on both equilibrium and dynamical properties. By developing methods for rigorously moving between the varying levels of detail in describing the molecules, all within a single integrated simulation package, the ambitious Doi Project in Japan, profiled in the site reports, seeks to develop a unified simulation capability for polymers.

Beyond bead-spring models, even more coarse-grained descriptions exist, such as dissipative particle dynamics (DPD) (Bonet Avalos and Mackie 1999; Evans 1999; Flekkøy and Coveney 1999; Groot, Madden, and Tildesley 1999; Groot and Warren 1997; Marsh and Coveney 1998; Schlijper, Hoogerbrugge, and Manke 1995). The connection between DPD and the macroscopic (hydrodynamic, Navier-Stokes) level of description has been made (Espanol 1995; Groot and Warren 1997), and very recently, the connection between DPD and atomistic molecular dynamics of spherically symmetric molecules has been developed (Marsh and Coveney 1998). However, the connection between a DPD representation of polymers and the underlying EA model is not rigorous. Likewise, time-dependent Landau-Ginzberg models (Glotzer 1995; Karim et al. 1998; Langer and Glotzer 1997) offer qualitative descriptions of phase separation in polymer blends but lack connection with an underlying EA model. Developing such connections will be a key to making these methods useful for product design.

The development of simulation methods for polymers is an active field that engages some of the most talented people in the molecular modeling field. It is an area of research that has significant intellectual challenges as well as significant practical application for the results of the research. Developing a hierarchy of coarse-grained models in which each member of the hierarchy can be rigorously related to the models both above and below it is the key to developing an integrated simulation strategy that meets the major molecular modeling needs in this area. We can expect to see progress in this area over the coming decade, with the most comprehensive and best-funded (64 man-years of effort) being the Doi Project described in the site reports (Appendix B.2).

ACKNOWLEDGEMENT

We acknowledge very helpful discussions with Keith Gubbins, who particularly influenced the choice of methods highlighted in the section entitled, "Molecular Simulation Methods."

REFERENCES

Allen, M.P., and D.J. Tildesley. 1987. *Computer Simulation of Liquids*. Oxford: Oxford University Press.

Andersen, H.C. 1980. Molecular-dynamics simulations at constant pressure and-or temperature. *J. Chem. Phys.* 72:2384-2393.

Barton, T.J., L.M. Bull, W.G. Klemperer, D.A. Loy, B. McEnaney, M. Misono, P.A. Monson, G. Pez, G.W. Scherer, J.C. Vartuli, and O.M. Yaghi. 1999. Tailored porous materials. *Chemistry of Materials* 11 (10):2633-2656.

Biggin, P.C., and M.S.P. Sansom. 1999. Interactions of alpha-helices with lipid bilayers: a review of simulation studies. *Biophysical Chemistry* 76 (3):161-183.

Binder, K. 1997. Applications of Monte Carlo methods to statistical physics. *Reports on Progress in Physics* 60 (5):487-559.

Binder, K., A. Baumgartner, A.N. Burkitt, D. Ceperley, A.M. Ferrenberg, D.W. Heermann, H.J. Herrmann, D.P. Landau, W. vonderLinden, H. DeRaedt, K.E. Schmidt, W. Selke, D. Stauffer, and A.P. Young. 1995. Recent developments in the Monte Carlo simulation of condensed matter. In *Monte Carlo Method in Condensed Matter Physics, Second, Corrected and Updated Edition*. ISBN 3-540-60174-0.

Binder, K., and W. Paul. 1997. Monte Carlo simulations of polymer dynamics: Recent advances. *Journal of Polymer Science Part B-Polymer Physics* 35 (1):1-31.

Bird, R.B., C.F. Curtiss, R.C. Armstrong, and O. Hassager. 1987. *Dynamics of Polymeric Liquids: Kinetic Theory.* Second ed. 2 vols. Vol. 2. New York: John Wiley and Sons.

Blochl, P.E., P. Margl, and K. Schwarz. 1996. Ab initio molecular dynamics with the projector augmented wave method. In B.B. Laird, R.B. Ross, and T. Ziegler (Eds.). *Chemical Applications of Density-Functional Theory.* Washington, DC. American Chemical Society. ACS Symposium Series 692.See http://www.oup-usa.org/toc/tc_0841234035.html and http://www.oup-usa.org/isbn/0841234035.html.

Boero, M., K. Terakura, T. Ikeshoji, C.C. Liew, and M. Parrinello. 2000. Car-Parrinello simulation of water at supercritical conditions. *Progress of Theoretical Physics Supplement* (138):259-261.

Bonet Avalos, J., and A.D. Mackie. 1999. Dynamic and transport properties of dissipative particle dynamics with energy conservation. *Journal of Chemical Physics* 111:5267-5276.

Bruge, F., M. Bernasconi, and M. Parrinello. 1999. Ab initio simulation of rotational dynamics of solvated ammonium ion in water. *Journal of the American Chemical Society* 121 (47):10883-10888.

Car, R., and M. Parrinello. 1985. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Letts.* 55:2471.

Chan, H.S., and K.A. Dill. 1993. The protein folding problem. *Physics Today* 46 (2):24-32.

Chen, B., and J.I. Siepmann. 1999. Transferable potentials for phase equilibria. 3. Explicit-hydrogen description of normal alkanes. *Journal of Physical Chemistry B* 103 (25):5370-5379.

Chialvo, A.A., and P.T. Cummings. 1999. Molecular-based modeling of water and aqueous solutions at supercritical conditions. In *Advances in Chemical Physics,* Vol. 109.

Cochran, H.D., P.T. Cummings, S.T. Cui, S.A. Gupta, R.K. Bhupathiraju, and P.F. LoCascio. 1998. Classical molecular simulations of complex, industrially-important systems on the Intel Paragon. *Computers and Mathematics with Applications* 35:73-84.

Cracknell, R.F., K.E. Gubbins, M. Maddox, and D. Nicholson. 1995. Modeling fluid behavior in well-characterized porous materials. *Accounts of Chemical Research* 28 (7):281-288.

Cui, S.T., P.T. Cummings, and H.D. Cochran. 1996. Multiple time step nonequilibrium molecular dynamics simulation of n-decane. *Journal of Chemical Physics* 104:255-262.

Cui, S.T., P.T. Cummings, H.D. Cochran, J.D. Moore, and S.A. Gupta. 1998. Nonequilibrium molecular dynamics simulation of the rheology of linear and branched alkanes. *International Journal of Thermophysics* 19 (2):449-459.

Cummings, P.T., and D.J. Evans. 1992. Molecular approaches to transport properties and non-Newtonian rheology. *Ind. Eng. Chem. Research* 31:1237-1252.

de Pablo, J., M. Laso, and U.W. Suter. 1992. Estimation of the chemical potential of chain molecules by simulation. *J. Chem. Phys.* 96:6157-6162.

de Pablo, J.J., and F.A. Escobedo. 1999. Monte Carlo methods for polymeric systems. In *Monte Carlo Methods in Chemical Physics.* eds. D.M. Ferguson, J.I. Siepman and D.G. Truhlar. Wiley.

de Pablo, J.J., M. Laso, J.I. Siepmann, and U.W. Suter. 1993. Continuum-configurational-bias Monte-Carlo simulations of long-chain alkanes. *Molecular Physics* 80 (1):55-63.

de Pablo, J.J., Q.L. Yan, and F.A. Escobedo. 1999. Simulation of phase transitions in fluids. *Annual Review of Physical Chemistry* 50:377-411.

Duan, Y., and P.A. Kollman. 1998. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 282 (5389):740-744.

Ermak, D.L., and J.A. McCammon. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 64:1352.

Errington, J.R., and A.Z. Panagiotopoulos. 1998. Phase equilibria of the modified Buckingham exponential-6 potential from Hamiltonian scaling grand canonical Monte Carlo. *J. Chem. Phys.* 109:1093-1100.

Escobedo, F.A., and J.J. de Pablo. 1999. Molecular simulation of polymeric networks and gels: Phase behavior and swelling. *Physics Reports-Review Section of Physics Letters* 318 (3):86-112.

Espanol, P. 1995. Hydrodynamics from dissipative particle dynamics. *Physical Review E* 52 (2):1734-1742.

Evans, D.J., W.G. Hoover, B.H. Failor, B. Moran, and A.J.C. Ladd. 1983. Nonequilibrium molecular dynamics via Gauss's principle of least constraint. *Phys. Rev. A* 28:1016.

Evans, D.J., and G.P. Morriss. 1990. *Statistical Mechanics of Nonequilibrium Liquids*. New York: Academic Press.

Evans, G.T. 1999. Dissipative particle dynamics: Transport coefficients. *Journal of Chemical Physics* 110:1338-1342.

Ferrenberg, A.M., and R.H. Swendsen. 1988. New Monte-Carlo technique for studying phase-transitions. *Physical Review Letters* 61 (23):2635-2638.

Ferrenberg, A.M., and R.H. Swendsen. 1989. Optimized Monte-Carlo data-analysis. *Physical Review Letters* 63 (12):1195-1198.

Flekkøy, E.G., and P.V. Coveney. 1999. From molecular dynamics to dissipative particle dynamics. *Physical Review Letters* 83:1775-1778.

Freire, J.J. 1999. Conformational properties of branched polymers: Theory and simulations. In *Branched Polymers Ii*.

Frenkel, D., and B. Smit. 1996. *Understanding Molecular Simulation: From Algorithms to Applications*. San Diego: Academic Press.

Geissler, P.L., C. Dellago, D. Chandler, J. Hutter, and M. Parrinello. 2000. Ab initio analysis of proton transfer dynamics in $(H_2O)_3H^+$. *Chemical Physics Letters* 321 (3-4):225-230.

Gelb, L.D., K.E. Gubbins, R. Radhakrishnan, and M. Sliwinska-Bartkowiak. 1999. Phase separation in confined systems. *Reports on Progress in Physics* 62 (12):1573-1659.

Gillespie, D.T. 1968. Ph.D. dissertation. Johns Hopkins University.

Gillespie, D.T. 1976. General method for numerically simulating stochastic time evolution of coupled chemical-reactions. *Journal of Computational Physics* 22 (4):403-434.

Glotzer, S.C. 1995. Computer simulations of spinodal decomposition in polymer blends. In *Annual Reviews of Computational Physics*, edited by D. Stauffer. Singapore: World Scientific.

Groot, R.D., T.J. Madden, and D.J. Tildesley. 1999. On the role of hydrodynamic interactions in block copolymer microphase separation. *Journal of Chemical Physics* 110:9739-9749.

Groot, R.D., and P.B. Warren. 1997. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *Journal of Chemical Physics* 107:4423-4435.

Haliloglu, T., I. Bahar, B. Erman, E.G. Kim, and W.L. Mattice. 1996. A dynamic rotational isomeric state approach for extension of the time scale of the local dynamics observed in fully atomistic molecular dynamics simulations: Application to polybutadiene. *Journal of Chemical Physics* 104 (12):4828-4834.

Haliloglu, T., and W.L. Mattice. 1998. Mapping of rotational isomeric state chains with asymmetric torsional potential energy functions on a high coordination lattice: Application to polypropylene. *J. Chem. Phys.* 108:6989-6995.

Hofmann, D., L. Fritz, J. Ulbrich, C. Schepers, and M. Bohning. 2000. Detailed-atomistic molecular modeling of small molecule diffusion and solution processes in polymeric membrane materials. *Macromolecular Theory and Simulations* 9 (6):293-327.

Johnson, J.K., A.Z. Panagiotopoulos, and K.E. Gubbins. 1994. Reactive canonical Monte-Carlo—A new simulation technique for reacting or associating fluids. *Molecular Physics* 81 (3):717-733.

Karim, A., J.F. Douglas, B.P. Lee, S.C. Glotzer, J.A. Rogers, R.J. Jackman, E.J. Amis, and G.M. Whitesides. 1998. Phase separation of ultrathin polymer-blend films on patterned substrates. *Phys. Rev. E* 57:R6273.

Kofke, D.A. 1993. Gibbs-Duhem Integration: A new method for direct evaluation of phase coexistence by molecular simulation (preliminary communication). *Molec. Phys.* 78:1331-1336.

Kremer, K., and G.S. Grest. 1990. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* 92:5057-5086.

Ladd, A.J.C. 1984. Equations of motion for non-equilibrium molecular dynamics simulations of viscous flow in molecular fluids. *Molecular Physics* 53 (2):459-463.

Langer, S.A., and S.C. Glotzer. 1997. Morphogenesis in polymer/nematic liquid crystal blends. *Physica A* 239:358.

Laso, M., J.J. de Pablo, and U.W. Suter. 1992. Simulation of phase-equilibria for chain molecules. *J. Chem. Phys.* 97:2817-2819.

Lees, A.W., and S.F. Edwards. 1972. The computer study of transport processes under extreme conditions. *J. Phys. C: Solid State* 5:1921.

Levi, A.C., and M. Kotrla. 1997. Theory and simulation of crystal growth. *Journal of Physics-Condensed Matter* 9 (2):299-344.

MacKerell, A.D., D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* 102 (18):3586-3616.

Marsh, C.A., and P.V. Coveney. 1998. Detailed balance and H-theorems for dissipative particle dynamics. *Journal of Physics A* 31:6561-8.

Martin, G. 1998. Modelling materials driven far from equilibrium. *Current Opinion in Solid State & Materials Science* 3 (6):552-557.

Martin, M.G., and J.I. Siepmann. 1998. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J. Phys. Chem.* 102:2569-77.

Marx, D., M. Sprik, and M. Parrinello. 1997. Ab initio molecular dynamics of ion solvation. The case of Be2+ in water. *Chemical Physics Letters* 273 (5-6):360-366.

Marx, D., M.E. Tuckerman, J. Hutter, and M. Parrinello. 1999. The nature of the hydrated excess proton in water. *Nature* 397 (6720):601-604.

Marx, D., M.E. Tuckerman, and M. Parrinello. 2000. Solvated excess protons in water: quantum effects on the hydration structure. *Journal of Physics-Condensed Matter* 12 (8A):A153-A159.

McCabe, C., S.T. Cui, P.T. Cummings, P.A. Gordon, and R.B. Saeger. 2001. Examining the rheology of 9-octylheptadecane to giga-pascal pressures. *J. Chem. Phys.* 114. 1887-1891.

McCabe, C., P.T. Cummings, and S.T. Cui. 2000. Characterizing the viscosity-temperature dependence of lubricants by molecular simulation. *Fluid Phase Equil.* submitted for publication.

McQuarrie, D.A. 1976. *Statistical Mechanics*: Harper and Row, New York.

Mehta, M., and D.A. Kofke. 1993. Implementation of the Gibbs ensemble using a thermodynamic model for one of the coexisting phases. *Molec. Phys.* 79:39-52.

Mooij, G.C.A.M., D. Frenkel, and B. Smit. 1992. Direct simulation of phase-equilibria of chain molecules. *J. Phys. Condens. Matter* 4:L255-L259.

Moore, J.D., S. Cui, H.D. Cochran, and P.T. Cummings. 2000a. Molecular dynamics study of a short-chain polyethylene melt. I. Steady-state shear. *Journal of Non-Newtonian Fluid Mechanics* 93:83-99.

Moore, J.D., S. Cui, H.D. Cochran, and P.T. Cummings. 2000b. Molecular dynamics study of a short-chain polyethylene melt. II. Transient response upon onset of shear. *Journal of Non-Newtonian Fluid Mechanics* 93:101-116.

Moore, J.D., S.T. Cui, H.D. Cochran, and P.T. Cummings. 2000c. Rheology of lubricant basestocks: A molecular dynamics study of C_{30} isomers. *J. Chem. Phys.* 113 (19):8833-8840.

Moore, J.D., S.T. Cui, H.D. Cochran, and P.T. Cummings. 1999. The transient rheology of a polyethylene melt under shear. *Physical Review E* 60:6956-6959.

Moore, J.D., S.T. Cui, P.T. Cummings, and H.D. Cochran. 1997. Lubricant characterization by molecular simulation. *A.I.Ch.E. Journal* 43:3260-3263.

Mundy, C.J., J. Hutter, and M. Parrinello. 2000. Microsolvation and chemical reactivity of sodium and water clusters. *Journal of the American Chemical Society* 122 (19):4837-4838.

Nakano, A., R.K. Kalia, and P. Vashishta. 1999. Scalable molecular-dynamics, visualization, and data-management algorithms for materials simulations. *Computing in Science & Engineering* 1 (5):39-47.

Nath, S.K., F.A. Escobedo, and J.J. de Pablo. 1998. On the simulation of vapor-liquid equilibria for alkanes. *Journal of Chemical Physics* 108 (23):9905-9911.

Neumaier, A. 1997. Molecular modeling of proteins and mathematical prediction of protein structure. *Siam Review* 39 (3):407-460.

Norman, G.E., and V.S. Filinov. 1969. Investigations of phase transitions by a Monte Carlo method. *High Temp.* (USSR) 7:216-22.

Onuchic, J.N., H. Nymeyer, A.E. Garcia, J. Chahine, and N.D. Socci. 2000. The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. In *Advances in Protein Chemistry*. Vol 53. F.M. Richards, D. Eisenberg, P.S. Kim, C.R. Matthews (eds.). Academic Press. 88-153.

Orkoulas, G., and A.Z. Panagiotopoulos. 1999. Phase behavior of the restricted primitive model and square-well fluids from Monte Carlo simulations in the grand canonical ensemble. *J. Chem. Phys.* 110:1581-1590.

Panagiotopoulos, A.Z. 1987a. Adsorption and capillary condensation of fluids in cylindrical pores by Monte Carlo simulation in the Gibbs ensemble. *Molec. Phys.* 62:701.

Panagiotopoulos, A.Z. 1987b. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Molec. Phys.* 61:813.

Panagiotopoulos, A.Z. 1992. Direct determination of fluid-phase equilibria by simulation in the Gibbs ensemble—A review. *Molecular Simulation* 9 (1):1-23.

Panagiotopoulos, A.Z. 1996. Current advances in Monte Carlo methods. *Fluid Phase Equilibria* 116 (1-2):257-266.

Panagiotopoulos, A.Z. 2000. Monte Carlo methods for phase equilibria of fluids. *Journal of Physics-Condensed Matter* 12 (3):R25-R52.

Panagiotopoulos, A.Z., N. Quirke, M. Stapleton, and D.J. Tildesley. 1988. Phase equilibria by simulation in the Gibbs ensemble. Alternative derivation, generation, and application to mixture and membrane equilibria. *Mol. Phys.* 63:527.

Pande, V.S., A.Y. Grosberg, and T. Tanaka. 2000. Heteropolymer freezing and design: Towards physical models of protein folding. *Reviews of Modern Physics* 72 (1):259-314.

Parrinello, M. 1997. From silicon to RNA: The coming of age of ab initio molecular dynamics. *Solid State Communications* 102 (2-3):107-120.

Phung, T.N., J.F. Brady, and G. Bossis. 1996. Stokesian dynamics simulation of Brownian suspensions. *Journal of Fluid Mechanics* 313:181-207.

Pinches, M.R.S., D.J. Tildesley, and W. Smith. 1991. Large scale molecular dynamics on parallel computers using the link-cell algorithm. *Molecular Simulation* 6:51-87.

Plimpton, S. 1995. Fast parallel algorithms for short-range molecular-dynamics. *Journal of Computational Physics* 117 (1):1-19.

Plimpton, S., and B. Hendrickson. 1995. Parallel molecular-dynamics algorithms for simulation of molecular-systems. In *Parallel Computing in Computational Chemistry*. T.G. Mattson (ed.). ACS Symposium Series 592. 114-132.

Plimpton, S., and B. Hendrickson. 1996. A new parallel method for molecular dynamics simulation of macromolecular systems. *Journal of Computational Chemistry* 17 (3):326-337.

Radeke, M.R., and E.A. Carter. 1997. Ab initio dynamics of surface chemistry. *Annual Review of Physical Chemistry* 48:243-270.

Ramaniah, L.M., M. Bernasconi, and M. Parrinello. 1998. Density-functional study of hydration of sodium in water clusters. *Journal of Chemical Physics* 109 (16):6839-6843.

Rognan, D. 1998. Molecular dynamics simulations: A tool for drug design. *Perspectives in Drug Discovery and Design* 9-11:181-209.

Rudisill, J.W., and P.T. Cummings. 1991. The contribution of internal degrees of freedom to the non-Newtonian rheology of model polymer fluids. *Rheologica Acta* 30:33-43.

Sarman, S., D.J. Evans, and P.T. Cummings. 1998. Recent developments in non-equilibrium molecular dynamics. *Physics Reports* 305:1-92.

Schlijper, A.G., P.J. Hoogerbrugge, and C.W. Manke. 1995. Computer simulation of dilute polymer solutions with the dissipative particle dynamics method. *Journal of Rheology* 39 (3):567-579.

Silvestrelli, P.L., and M. Parrinello. 1999. Structural, electronic, and bonding properties of liquid water from first principles. *Journal of Chemical Physics* 111 (8):3572-3580.

Smith, W.R., and B. Triska. 1994. The reaction ensemble method for the computer-simulation of chemical and phase-equilibria .1. Theory and basic examples. *Journal of Chemical Physics* 100 (4):3019-3027.

Sorensen, M.R., and A.F. Voter. 2000. Temperature-accelerated dynamics for simulation of infrequent events. *Journal of Chemical Physics* 112 (21):9599-9606.

Sprik, M., J. Hutter, and M. Parrinello. 1996. Ab initio molecular dynamics simulation of liquid water: Comparison three gradient-corrected density functionals. *Journal of Chemical Physics* 105 (3):1142-1152.

Stoltze, P. 2000. Microkinetic simulation of catalytic reactions. *Progress in Surface Science* 65 (3-4):65-150.

Swendsen, R.H., J.S. Wang, and A.M. Ferrenberg. 1995. New Monte Carlo methods for improved efficiency of computer simulations in statistical mechanics. In *Monte Carlo Method in Condensed Matter Physics, Second, Corrected and Updated Edition. Topics in Applied Physics*. vol 71. Springer-Verlag.

Trout, B.L., and M. Parrinello. 1998. The dissociation mechanism of H_2O in water studied by first-principles molecular dynamics. *Chemical Physics Letters* 288 (2-4):343-347.

Trout, B.L., and M. Parrinello. 1999. Analysis of the dissociation of H_2O in water using first-principles molecular dynamics. *Journal of Physical Chemistry B* 103 (34):7340-7345.

Tuckerman, M.E., B.J. Berne, and G.J. Martyna. 1992. Reversible multiple time step molecular dynamics. *J. Chem. Phys.* 97:1990.

Tuckerman, M.E., and G.J. Martyna. 2000. Understanding modern molecular dynamics: Techniques and applications. *Journal of Physical Chemistry B* 104 (2):159-178.

Uhlherr, A., and D.N. Theodorou. 1998. Hierarchical simulation approach to structure and dynamics of polymers. *Current Opinion in Solid State & Materials Science* 3 (6):544-551.

Voter, A.F. 1997a. Accelerating the dynamics of infrequent events. *Abstracts of Papers of the American Chemical Society* 213:284-PHYS.

Voter, A.F. 1997b. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Physical Review Letters* 78 (20):3908-3911.

Voter, A.F. 1997c. A method for accelerating the molecular dynamics simulation of infrequent events. *Journal of Chemical Physics* 106 (11):4665-4677.

Voter, A.F. 1998. Parallel replica method for dynamics of infrequent events. *Physical Review B-Condensed Matter* 57 (22):R13985-R13988.

Wick, C.D., M.G. Martin, and J.I. Siepmann. 2000. Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and alkylbenzenes. *Journal of Physical Chemistry B* 104 (33):8008-8016.

Zachariah, M.R., and M.J. Carrier. 1999. Molecular dynamics computation of gas-phase nanoparticle sintering: A comparison with phenomenological models. *Journal of Aerosol Science* 30 (9):1139-1151.

Zwanzig, R. 1965. Time correlation functions and transport coefficients in statistical mechanics. *Ann. Rev. Phys. Chem.* 16:67.

CHAPTER 4

SCIENCE: INFORMATION TECHNOLOGIES

Peter A. Kollman

INTRODUCTION

In addition to quantum mechanical calculations and simulations, discussed in the previous two chapters, there have been other modeling approaches that have had equal or often greater impact in industrial applications. Unlike the above two methods, which evaluate energy and then use this to describe the thermodynamics of processes, these non-energy-based methods are generally much more qualitative in nature.

COMPUTER GRAPHICS

Improvement in computer graphics has been probably the single largest contributor to a wide use of molecularly based modeling. Although some specific history is mentioned below in the chapter on molecular modeling in the pharmaceutical industry, the ability to visualize the drug target (when its structure was known) or a conformational family of active and inactive analogs was essential in the involvement of non-modeling experts (i.e., the synthetic chemists) in the modeling. Computer graphics came of age when one could use depth cueing and stereo in color and thus highlight important interactions (Langridge et al. 1981). It has continued to develop and allow ever more beautiful and imaginative representations of important processes of chemical and biophysical interest (Ferrin and Klein 1998).

DISTANCE GEOMETRY

A second area of scientific development that has had an impact on modeling in biophysical chemistry and the pharmaceutical industry has been distance geometry. The transfer of this esoteric mathematical subject to chemistry and molecular systems was done by Crippen, Havel and Kuntz in the late 1970s (Crippen and Havel 1988) and used both in NMR structure determination and in QSAR of ligand binding. Crippen has continued to develop it for modeling ligand-macromolecule interactions (Crippen 1999).

QSAR/QSPR

A third area of modeling involves correlations of experimental results with some physical or non-physical property of the system. These QSAR/QSPR approaches have a long history in modeling of biological systems and are now very important for analyzing and using the results of molecular modeling in the chemical industries.

A typical situation for using these techniques is this: Given that one has some physical property that one wants to optimize (drug efficacy, enzyme binding, tensile strength of a polymer), how can one approach this problem? Among the earliest approaches to this problem were those of Hansch (1971), who related

biological activity of a molecule to its hydrophobicity, measured as its octanol/water partition coefficient. This approach is along the lines of linear free energy relationships in physical organic chemistry (e.g., the Hammett relationship (Hammett 1935)), where the free energy (e.g., logarithm of an equilibrium constant) is related to another energy- or free-energy-based property. As suggested, the simplest such relationships are developed as linear models, and linear regression is used. As the model gets more complex, one often turns to non-linear regression (Kowalski 1984) and can use neural networks (Andrea and Kalayeh 1991) to "train" such models—which relate some "activity" to some set of physical properties.

Besides hydrophobicity, other terms of possible relevance as independent variables in such models are steric/van der Waals and electrostatic properties. A method that takes three-dimensional information into account is the COMFA (comparative molecular field analysis) approach (Cramer, Patterson, and Bunce 1988). It and its variants have been very popular and useful in optimizing ligand binding to macromolecules, taking into account the three-dimensional properties (i.e., their electrostatic and van der Waals potentials surrounding them).

The above techniques are called QSAR (quantitative structure-activity relationships). A related approach is QSPR (quantitative structure-properties relationships), where the optimization of some physical property other than activity is the goal (Katrizky, Lobanov, and Karelson 1995). Once again, one can do this in the context of various linear and non-linear models and can use various properties of the collection of molecules in such a model. In order to derive the electrostatic properties or electronic properties, to test if they are of relevance in such relationships, quantum mechanical calculations can be used. Since the ultimate goal is one where the precise electronic structure is not required, less accurate and more efficient (e.g., semi-empirical) quantum mechanical methods may be adequate for such purposes.

USING SIMULATIONS IN STRUCTURE DETERMINATION AND REFINEMENT

Although molecular dynamics per se is discussed in Chapter 3, the use of molecular dynamics in X-ray structure refinement and NMR structure determination is one of the great uses and successes of molecular modeling. It has been crucial both for protein crystallography and for establishing the structures of new zeolites. Since the X-ray data in macromolecular crystallography are insufficient to fully define the structure at atomic resolution, in contrast to the typical case for a small-molecule crystal, one needs to use model-building to fill in the gaps.

The XPLOR program, developed by Brunger and Karplus (Brunger, Kuriyan, and Karplus 1987), is by far the most popular program for combined molecular dynamics/X-ray refinement and provides a good example of the general methodology. In it, one refines the initial structure using a function that contains both the difference between calculated and observed reflections and the molecular mechanical energy function. By using molecular dynamics and high temperature, one can move the structure out of local minimum and, by changing the weights of the molecular mechanical energy and the fit to experimental reflections, guide the system to a structure that has good stereochemistry and bonding and accurately reproduces the experimental reflections. (On our visits, one group leader of molecular modeling, not coincidentally a crystallographer by training, expressed frustration that molecular modeling in drug design had not had the impact or success of crystallographic refinement using molecular dynamics.)

Modeling is even more important in deriving macromolecular structures by NMR, where the number of experimental geometric data is significantly less than in crystallography (Case, Wiley, and Chichester 1998). Thus, using various technologies to guide the determination of the three-dimensional structure, such as distance geometry and molecular dynamics, plays a large role in NMR structure determination of macromolecules.

DOCKING INTO BIOLOGICAL TARGETS

If one has the structure of a macromolecular target but not of its complex with ligands, it is essential that one should have computational methods that can accurately predict the structure of the ligand protein complex, not only for a single ligand but for an entire database of ligands. The programs DOCK (Zou, Sun, and Kuntz

1999) and FLEXX (Kramer et al. 1999) enable one to screen databases of the order of 100,000 ligands and screen for those most likely to bind to a target macromolecule. At the other end of the spectrum is the program AUTODOCK (Rao and Olsen 1999), which carries out more extensive and lengthy docking of a single ligand using a Monte-Carlo-based method. Methods to carry out docking for entire combinatorial libraries are being developed (Kick et al. 1997) and should have a major impact in structure-based ligand design.

BIOINFORMATICS AND CHEMICAL INFORMATICS

Being able to analyze, manipulate and query the entire genome sequences of various organisms is the target of bioinformatics; carrying out the same kind of analysis on large databases of organic molecules, including combinatorial libraries, is the target of cheminformatics.

The goal of bioinformatics is to combine gene discovery, gene function, drug design, and drug development. The significant commercial potential of bioinformatics is highlighted in a recent article (Thayer 2000). Within the table on p. 20 of that article is the focus of cheminformatics as well. Thus, one might view both chem- and bioinformatics as part of a larger effort to catalog, store, manipulate, and analyze the vast data on genes, gene expression and the expressed proteins, and to characterize the effect that small molecules can have on DNA, RNA, and protein targets. Both chem- and bioinformatics are large and important disciplines, and many of the pharmaceutical companies we visited had significant efforts and often separate groups in these areas. One group leader of an American company noted his desire that these approaches become more fully integrated into and part of what is considered "computational chemistry and modeling," to the benefit of both computational chemistry and informatics.

These activities point toward the future use of similar methods in the chemical and materials industries. Rzepa (see site report for Royal Society of Chemistry, UK) and others have noted that information analysis and visualization technologies are just beginning to be put to use in such ways.

COMPUTER SCIENCE

Many of the calculations carried out in computational chemistry require powerful computational resources, and often the full harnessing of these resources requires developments in computer science. The same is true for information analysis. The software for vector and parallel computing has had great influence on computational chemistry over the last 15 years, as has the hardware development of RISC processors and, more recently, Beowulf clusters. It seems likely that the continued increase in capabilities of computer hardware and the harnessing of it by new software will have great impacts on computational chemistry and molecular modeling both in academia and industry. These issues are more fully discussed in the chapter on infrastructure (Chapter 9).

REFERENCES

Andrea, T., and H. Kalayeh. 1991. *J. Med. Chem.* v. 34. p. 2824.

Brunger, A.T., J. Kuriyan, and M. Karplus. 1987. *Science.* 235, 458.

Case, D., J. Wiley, and Chichester. 1998. NMR refinement. *Encyclopedia of Computational Chemistry.* P.v.R. Schleyer (ed.). p. 1866.

Cramer, R.D., D.E. Patterson, and J.D. Bunce. 1988. *J. Amer. Chem. Soc.* v. 110. p. 5959.

Crippen, G. 1999. *J. Comp. Chem.* v. 20. p. 1577.

Crippen, G.M., and T.F. Havel. 1988. *Distance Geometry and Molecular Conformations.* Research Studies Press. Wiley. New York.

Ferrin, T.E., and T.E. Klein. 1998. Computer graphics and molecular modeling. *Encyclopedia of Computational Chemistry.* P.v.R. Schleyer (ed.). J. Wiley. Chichester. p. 463.

Hammett, L.P. 1935. *Chem. Rev.* v. 17. 125.

Hansch, C. 1971. *Drug Design*, E.J. Ariens (ed.). Academic Press. New York. vol. 1, chapter 2.

Katrizky, A., V. Lobanov, and M. Karelson. 1995. *Chem. Soc. Rev.* v. 24. 279.

Kowalski, B. (ed.). 1984. *Chemometrics—Mathematics and Statistics in Chemistry*. D. Reidel Publishing Co. Dordecht.

Joliffe, I.T. 1986. *Principal Component Analysis*. Springer-Verlag. New York.

Kramer, B., G. Metz, M. Rarey, and T. Lengauer. 1999. *Med. Chem. Res.* 9. 463.

Kick, E.M., E.M.D.C. Roe, A.G. Skillman, G.C. Liu, T.J.A. Ewing, Y.X. Sun, I.D. Kuntz, and J.A. Ellman. 1997. *Chemistry and Biology*. 4. 297.

Langridge, R., T.E. Ferrin, I.D. Kuntz, and M.L. Connolly. 1981. *Science*. 211. p. 611.

Rao, M.S., and A.J. Olson. 1999. *Proteins—Structure, Function and Genetics* .34. 173.

Zou, X.Q., Y.X. Sun, and I.D. Kuntz. 1999. *J. Amer. Chem. Soc.* 121. 8033.

Thayer, A.M. 2000. Bioinformatics for the masses. *Chem. Eng. News.* 19.

CHAPTER 5

APPLICATIONS: THE CHEMICAL INDUSTRY

Anne M. Chaka

INTRODUCTION

The chemical industry is among the United States' most successful industries, with sales of $419 billion per year. In 1998, U.S. chemical exports totaled over $68 billion, larger than either agriculture or aircraft/aerospace, making it the largest exporting sector. More than 10% of all U.S. exports are chemicals and related products. Taking into account imports of nearly $55 billion, this resulted in a 1998 trade surplus of over $13 billion. The chemical industry has been recording trade surpluses for more than 70 straight years, making it the country's premier exporting industry.

The continuing success of the chemical industry is a result of its ability to evolve in response to new technological and societal needs. For example, among industrial and manufacturing groups, it is the leading investor in research and development. Over $26 billion is expended annually in research and development, funding the efforts of 95,000 R&D scientists, engineers, and technicians, and accounting for one out of every seven patents issued annually in the United States.

Driven by the decades-long sustained exponential growth in computing power, molecular simulation and computational quantum chemistry are making significant inroads into the chemical and related industries, and they will increasingly do so in the coming decade. Indeed, these techniques, under the general title of computational chemistry, have been identified as a key technology for enabling the U.S. chemical industry to achieve its "vision" in 2020 (Oleson 1996). This vision includes using computational methodologies to design new products and energy efficient, environmentally benign processes to manufacture them.

This sector includes any industry that produces chemical products or components, such as consumer products, specialty chemicals, fuels and petroleum products, organic materials, and polymers. It includes companies described as being "chemical manufacturing industries" and the "chemical process industries (CPI)." Also included are relevant government laboratories and commercial software companies that conduct research of interest to the chemical industry. Excluded from the discussion in this chapter are pharmaceuticals, materials that are primarily inorganic, and the catalytic processes used in chemical manufacturing, as these areas are covered in Chapters 6, 7, and 8, respectively.

The set of companies and national labs included in the WTEC study cannot be comprehensive, but it is representative of broad sectors of the industry. Examples of companies included in this sector are shown in Table 5.1.

Table 5.1
Examples of Companies Involved in the Chemical Industries

Air Liquide	Air Products and Chemicals
Atofina (TotalFinaElf)	BASF
Bayer AG	BP
Dow Chemical	Dow-Corning
DSM	DuPont
Ford Motor Co.	Fujitsu
Institut Français du Petrole	JRCAT
Lubrizol	Mitsubishi
Rhodia	RIKEN
Rohm & Haas	Sumitomo
3M	Unilever

CHEMICAL INDUSTRY TRENDS AND DRIVERS: IMPACT OF AND ON MODELING

External economic and environmental pressures have a dramatic effect on the chemical industry, which in turn has an impact on how molecular and materials modeling is applied and supported within a company. Thirty years ago, there was a shortage of chemicals, and customers would pay premium prices for any chemical that worked at all. Trial and error was used with success to develop new chemistry. Today, however, the trend has shifted due to increased competition from an abundance of chemicals, customer consolidation, and global competition that is driving commodity pricing even for high-performance and fine chemicals. In addition, the chemicals produced today must be manufactured in a way that minimizes pollution and energy costs. Trial and error has become too costly, and the probability of success too low, for a company to rely heavily on an Edisonian approach to developing new products and manufacturing processes.

Today it is becoming widely recognized that companies need to develop and fine-tune chemicals and formulations by design in order to remain competitive, whether it is improving the skin feel of a cosmetic, the performance characteristics of a polymer, or the viscosity index of a lubricant. It is more effective to screen chemicals prior to a long and costly synthesis and testing process. Manufacturing has becoming more streamlined and targeted towards producing the optimum performing structure for the lowest possible price with a minimal waste stream. Shifting from batch processing to continuous feedstreams necessitates a greater understanding of the reaction kinetics, and hence the mechanism, to optimize feedstream rates. Modeling and simulation was well-established in the drug industry over ten years ago, but only during the past five years has it begun to be firmly established in the chemical industry.

The need for global competitiveness and cost reduction has spawned a record number of mergers, acquisitions, and joint ventures in the petrochemical industry. In the past 10 years, six of the top ten olefin producers were formed from the merging of 14 companies (C&EN 2000), e.g., BP-Amoco (already changed to "BP"), ExxonMobil, and TotalFinaElf (Total, Petrofina, and Elf Aquitaine). Dow Chemical and Union Carbide merged in 2001. Degussa-Hüls, itself a merger, joined with SKW to become Degussa because of the merger of their parent companies Veba and Viag, and Rhône Poulenc and Hoechst underwent a complicated merger and rearrangement into Aventis, Rhodia, and Celanese AG. In most instances the modeling efforts survived these mergers. In companies like ICI and Eastman, molecular modeling has not survived company restructuring.

In general, the support for modeling programs within a company parallels the support for research. Not all sectors of the chemical industry (fine, specialty, commodity) are the same, nor are all companies within a sector the same. In general, longer time cycles are allowed for product design than for fire-fighting. Fundamental research is becoming more and more difficult to justify in the U.S. business climate. (This is much less true in Japan and somewhat less in Europe.) The chemical industry today will not support

fundamental work to understand chemistry that already works well enough. U.S. industry *will*, however, support fundamental work that does the following:

- Reduces time, waste, and cost to manufacture a chemical
- Improves quality control
- Reduces time and cost to develop new chemistry

The shift in emphasis away from fundamental research for its own sake is particularly apparent in the petroleum industry, where the research strategy has moved much more towards immediate business needs. Investment in modeling usually follows the investment in R&D. Many companies that once had strong modeling efforts have cut these efforts as part of significant overall reductions in R&D expenditures.

WHAT IS THE ROLE OF MOLECULAR AND MATERIALS MODELING IN THE CHEMICAL INDUSTRY?

The principal role of molecular and materials modeling in the chemical industry is to speed product development and guide experiment. Most of the modelers reported that one of their key functions involved developing mechanistic hypotheses and a rational, strategic approach to problem solving. This is most effectively done in close cooperation with experimentalists at the early stages of a project. The hypotheses are then tested either experimentally or computationally, depending on which method is most appropriate for the given question.

Most modeling groups reported that the most important deliverables from their projects are to provide understanding and design rules to enable better technical decision-making and ranking of potential chemical structures and formulations. If these design rules can be incorporated into a QSPR/QSAR (quantitative structure property/activity relationship) model, which could be queried by formulators in many areas across a company, this is seen as measurable success. It is a way of capturing corporate knowledge in a usable form and maximizing the utility of data.

Another of the most important roles of modeling has been in the elimination of dead-ends, to show that a product or process is not physically or thermodynamically possible.

Modeling is also highly valued because it can stimulate new ways of thinking. This is more difficult to quantify, but the companies we visited universally felt that graphics and visualization are invaluable for the communication of highly technical or difficult concepts to customers, coworkers, and management, and to enable creative thinking.

Virtually all of the chemical companies in the study used modeling and simulation to streamline process development in manufacturing, as well as the discovery process in research. This is particularly true for the commodity producers and petrochemical companies, where profit is largely determined by manufacturing costs. Modeling is used to improve the yield, selectivity, or waste reduction of a process, as well as to provide accurate thermochemical and kinetic data. In the early stages of process design, so-called process screening (Douglas 1988), the most commonly encountered problem is the lack of experimental data on systems of interest. In the process screening stage when process alternatives are being eliminated on the basis of gross cost differences, the process designer will typically employ engineering correlations or commission crude, low-cost experiments (such as chromatographic experiments to determine infinite dilution activity coefficients rather than a full vapor-liquid equilibrium experiment). It is in these early stages of process design that molecular modeling can have a significant impact by providing a route to thermophysical and thermochemical properties that is more fundamentally sound than correlation.

Lastly, modeling has begun to play a greater and greater role in the broadening of patent protection, particularly in the U.S. and Japan.

NEEDS OF THE CHEMICAL MANUFACTURING INDUSTRY

The vast majority of industrial problems are exceedingly complex and require a broad range of modeling and simulation tools, as well as a broad range of expertise from the modeling group. This can be particularly challenging if there are only one or two modelers in a company.

An example of this complexity can be found in the oil additive field. Performance of an antiwear chemical additive in motor oil, for example, is determined by a broad range of factors. Each of these factors requires a different methodology to evaluate through modeling and simulation, as is indicated in parentheses below:

- First, the additive must be miscible in the oil, yet it must partition to the surface in order to impart antiwear performance (statistical mechanics and/or empirical correlation methods).
- Then the additive must have sufficient thermal and oxidative stability (high-level quantum mechanics, which can include multi-reference techniques).
- The additive must not have unfavorable chemical reactions with other species in the formulation (QM).
- At the surface, the additive reacts with the metal or metal oxide surface (solid-state physics).
- Surface structures develop which are transformed under high temperatures and pressures to form a solid film tens to thousands of angstroms thick. These films have very different wear and mechanical properties depending upon their structure and composition (atomistic solid-state physics linked to coarse-grained material science and even finite-element methods).

This complexity of industrial problems, and the range of techniques and expertise required to address them, underscores the need for validation of methodologies across the board.

Such industrial problem-solving requires a different approach than that taken in academia, and it is important to understand these differences in order to develop tools that meet the needs of industry. Table 5.2 highlights some of the key differences. The approaches need not and should not be the same. Nevertheless, new graduates must be prepared to adapt to these approaches, and faculty should appreciate the differences.

Table 5.2
Comparison of Necessary Problem-Solving Skills in Industry and Academia

Industry	Academia
Complexity of entire problem must be treated	Can focus on one aspect of a problem to match tools and expertise
Multidisciplinary approach, usually working in teams	Usually work individually
Proprietary	Public
Often sufficient to solve 80% of a problem to eliminate dead ends	High accuracy and thorough understanding required for publication and advancing science
Following 'line of least resistance' often means making approximations on a difficult problem	Following 'line of least resistance' can mean avoiding a problem that is too difficult to treat accurately
Short time scales required for solving problems	Longer time scales
Cannot spend a full day to prepare input files for one calculation	Input files often prepared by hand
Need for robust, validated code and methods	Debugging and validating code is critical for student/post doc's education and advancing the science
Commercial software interface is invaluable for preparation of input files and visualization and communication of results to colleagues and managers with diverse backgrounds	Communicate results with peers in the same field (Good visualization and communication is always important)

PRESENT STATE OF THE ART: METHODS

Software in Use

Most chemical companies that had made an investment in molecular modeling, irrespective of size, had the following modeling tools available:

- *Ab initio* molecular quantum mechanics (wavefunction and DFT)
- *Ab initio* solid-state quantum mechanics (DFT)
- Semi-empirical quantum mechanics (such as AM1, ZINDO)
- Molecular mechanics and dynamics (organic interaction potentials)

In the polymer industry, lattice, coarse-grain, and mesoscale models are standard. Surprisingly, QSAR/QSPR was found to be nearly as pervasive in the chemical industry as in pharmaceuticals, with applications ranging from the formulation of paints and soaps to predicting intake valve deposits in BMW engines.

Increasingly important—but not yet standard—tools are the QM/MM hybrids, *ab initio* dynamics, and organic-inorganic interaction potentials for molecular mechanics and dynamics.

The mesoscale methodologies such as dissipative particle dynamics (DPD) and Mesodyne are in the early development stages at only a few of the largest companies such as Unilever and BASF.

The vast majority of software used is from commercial sources such as Accelrys / Molecular Simulations, Inc. and Gaussian. Academic programs, such as the quantum codes GAMESS and VASP, are frequently used by individuals who already have an extensive background in the field. A big advantage of commercial code that is often cited is the unified suite of software with an excellent graphical user interface, plus robustness and technical support.

Accessible Properties

Table 5.3 summarizes of the properties accessible to molecular simulation and computational quantum chemistry, as identified in a recent National Science Foundation workshop report (Cummings et al. 1998). Many of these properties are useful in the process design context, but those that are routinely used are highlighted within the table. In the case of some of the properties, they provide input to correlations (such as critical properties or QSAR); in other cases the computations provide properties directly for use in process design calculations (such as mixture phase equilibria). As the reliability and accountability of molecular-based calculations grows (via accurate error estimates in predicted properties), their use in the chemical and materials industries will continue to accelerate. Already there are accounts in the published literature of cases where molecular modeling has been used to discriminate between differing experimental data (McCabe et al. 2001) and engineering correlations (Siepmann, Karabomi, and Smit 1993).

Method Development and Programming within Industry

Method development and programming is done in-house by only the largest companies with extensive modeling efforts. In addition, the problem has to be one that is basic to the specific industry and projected to provide a competitive advantage for an extensive period of time. Examples do exist, though: Dow-Corning invested in the development of siloxane force field parameters, Unilever in Dissipative Particle Dynamics, BASF in Mesodyne, Dow Chemical in QSPR for polymeric materials, and Fujitsu in ZINDO developments to predict nonlinear optical properties of molecules for liquid crystal displays. Japanese companies also have invested in development of new methodologies by having their scientists spend as much as four years on site at an academic or government research institution, such as the project led by Professor Doi.

Table 5.3
Properties Accessible to Molecular Modeling and Simulation Techniques
(properties that may be required in process design are highlighted)

Category	Property	Category	Property
Spectroscopic/ Chemical	Spectra: Vibrational Electronic Photoelectric Magnetic Magneto-optical Frequency-dependent dielectric response Band gaps Magnetic moments Dipole moments Spin resonance Reaction pathways Reaction rates Activation energies Transition states Reactivity (nucleophilic, electrophilic, radical susceptibility, oxidative and thermal stability) Acidity/basicity Molecular structure Electron density	Thermophysical	Intramolecular: Bond lengths, angles Conformational energies Site occupancy Persistence length Tacticity Steric specificity Docking Intermolecular: Structure factor Aggregation equilibrium structure of oriented molecular systems Crystal structure Normal boiling point Critical properties Densities Vapor pressure Viscosities Surface tension Mixture phase equilibria Diffusivities Acentric factor Solubility factor Molecular size and surface area Activity coefficient at infinite dilution Henry's law constant Change in volume on mixing Change in enthalpy on mixing Change in Gibbs free energy on mixing VLE, PVT of mixtures Second and third virial coefficients Crystal habit Polymer void volume Adsorption isotherms Hydrophobicities
Mechanical	Stress-strain relationships Elastic moduli Glass temperature Poisson's ratio Tensile strength Yield stress Viscoelasticity		

Source: Cummings et al. 1998

Regional Strengths in Methodology

Several regional strengths in methodologies are apparent, and in general they reflect where a particular methodology was most widely developed in academic or government institutions. Europe is particularly strong in *ab initio* dynamics, heterogeneous catalysis, and materials, as well as liquid phase mesoscale methods (Unilever, Bayer, BASF). The United States is strong in multi-reference methods and in applying quantum mechanics to chemical engineering problems. Japan has the greatest research effort in developing computer hardware specific for a given computational application, although IBM's Blue Gene supercomputer project, designed to fold proteins, is a significant advance.

Today's corporations are global entities, but their expertise is often a reflection of regional academic strengths of the areas where their corporate R&D activities are located. This is because the primary method of adoption of technology by a company is through exchange of personnel.

APPLICATIONS

Credibility Establishers

The earliest successes that established modeling's credibility with respective companies, and hence within the industry, included the areas of molecular thermochemistry (Dupont), homogeneous catalysis (polyolefins industry), QSAR/QSPR, spectroscopic interpretation, and mechanism elucidation and design rules. Examples of notable successes to date are:

- Rhône-Poulenc
  - Anti-scratch additive for polyurethane coatings (Flory χ-parameter and relative reactivities via QM)
  - Valuable negative result: Developing a material to compete with its competitor was not possible using nylon basis
- Lubrizol
  - QSPR model for gasoline additive formulation reduced testing costs by 1/3 for predicting intake valve deposits in BMW, Ford, Honda
- Dow
  - Each $\bullet H_f$ calculation saved $50,000 testing costs in 1996—over $100,000 in 2000
- Mitsubishi Chemicals
  - 5% of patents from their Yokohama facility involve some computational modeling (one modeler has 20-30 patents)
  - Polymer membrane discovery
  - New dyes

Credibility Destroyers

Over the past two decades, several companies began modeling efforts but discontinued them for a variety of reasons. In many cases, the reasons have been non-technical, usually change in business direction or elimination of research and development activities. Nevertheless, it is instructive to examine technical issues that have been important problems.

In the mid- to late-1980s, several companies began molecular modeling for heterogeneous catalysis, including *ab initio* dynamics. The computational power available at the time, however, was insufficient for these efforts to succeed in producing useful results in a reasonable timeframe. Some groups that attempted to write much of their own code in-house also did not succeed, as the development time proved too long. In other companies, modeling was not successful for reasons that included the following:

- Modeling had been initiated to solve one high-priority, very specific problem for which modeling was not appropriate.
- Methods were required that were not yet fully developed.
- Modelers did not have sufficient expertise in an area, or chemists who were not trained in modeling were expected to do it.
- Modelers were brought in at the last stage when everything else had failed and time, patience, and budget were in short supply.
- Modeling was oversold and expectations were too high or inappropriate—after the mid-1980s when computer graphics became much better, it was often because attractive visuals ("pretty pictures") left managers with exaggerated impressions of the power of modeling, which proved disastrous when modelers failed to correct those impressions.

HOW MODELING TECHNOLOGY IS TRANSFERRED TO INDUSTRY

Modeling and simulation technology is transferred to industry from academia and the national labs in a wide variety of ways. The method most often cited as being effective is the transfer of personnel—i.e., people who have received training in academia take a job in industry. New modeling techniques which are not "standard" are in general not adopted by industry; the expertise is not present in-house to use them, and there is insufficient validation for non-experts in a field to quickly get up to speed and learn the limitations. Therefore, for a method to be widely adopted in industry, it must be in a commercial or easy-to-use package and have extensive validation studies already available in the literature.

In cases where a company does not have the internal expertise or resources to develop a needed method, quite often they will work with academic collaborators or consortia (MSI, GdR 12090). In principle, the NSF GOALI program facilitates exchanges of people, although the paperwork and difficulty of moving people with families has limited the program's success. Industrial sabbaticals are sometimes allowed in Europe and the U.S. (Ford, Lubrizol, Xerox) but are rare. In contrast, they are quite common in Japan, where a long-term view is taken of employee development and training.

Universities and national labs also transfer their technology to industry through licensing of technology and small startups (IFP, Gaussian, Semichem), as well as sponsoring continuing education courses and visiting scientist programs.

To a lesser extent, a few of the large companies such as Dow and IBM offer industrial post-docs to train students in industrial applications.

CHALLENGES AND NEEDS

The top requirements in the chemical industry for improvements in molecular and materials modeling can be summarized as "bigger, better, faster," more extensive validation, and improvements at each scale and for coupled scales (multiscale techniques).

"Bigger" refers to being able to do larger systems of greater complexity, ranging from multireference wavefunction quantum mechanical methods involving complex transition metal species, to mesoscale behavior in multiple phase and composite systems. "Better" reflects the need for greater accuracy for activation energies and thermochemistry in condensed phases, and to weak van der Waals interactions, for example. "Faster" enables the simulation of rare event processes such as predicting the thermal and oxidative stability of a plastic under a wide variety of weathering, and the effect of a "dirt" particle on the fracture mechanics of a polymeric/inorganic filler composite.

More extensive validation is required because of the complexity and range of modeling and simulation techniques required to solve real-world industrial problems. A modeling expert trained in one field needs to have extensive validation studies available to critically evaluate what assumptions can and cannot be made in a new area, as well as what level of theory is appropriate for a given system. In addition to validation, it is essential to have tools that enable greater understanding and interpretation of the results of modeling and simulation for *non-experts* in a given field as well as for the modeling experts. Non-experts can fruitfully use modeling for insights and ideas, to communicate their results to their coworkers and customers, and to maximize the interpretive value.

The rest of this section will elaborate on aspects of these requirements.

Greater Numerical Accuracy for Thermochemistry

This is critical for many challenges faced by the chemical industry. For example, a 1 kcal/mol difference in free energy of formation may result in a product ratio change of approximately 9:1. Gas-phase thermochemistry for molecules smaller than 15 heteroatoms is well established and considered routine. There is more uncertainty for radicals, excited states, and transition states. Furthermore, condensed-phase

thermochemistry is not yet sufficiently accurate for industrial applications, particularly if the condensed phase is a complex mixture. Greater accuracy for transition metals is also of particular importance.

Greater Accuracy for Kinetics

This is even more demanding than for thermochemistry, as rates depend exponentially on the activation energy. This is extremely important for manufacturing and process design, to optimize feed-stream rates, and to determine whether a process is kinetically or thermodynamically controlled. The energy of the activation barrier depends on the difference between structures in their equilibrium ground state(s) and the transition state where bonds are partially made or broken. Therefore correlation errors which often fortuitously cancel for thermochemistry often do *not* cancel for activation barriers. This problem is particularly serious for free-radical processes and transition metal catalysts.

More Extensive Reference Data

The demand for accurate kinetics has resulted in a clear list of needs for the chemical manufacturing industry. Development of accurate methods for calculating the energy of transition states has been hindered by the lack of a "G2-type data set" for activation barriers, e.g., an extensive, refereed and highly accurate set of kinetic data. This is crucial for many classes of organic reactions, but it is especially critical for free-radical processes, transition-metal centers, and inorganic species.

DFT and HF/DFT-hybrid Methods

These have been widely used for many problems in industry because they include electron correlation at a lower cost than post-Hartree-Fock methods. DFT-based methods, however, are not appropriate for all types of reactions. For example, in the $Cl\bullet + CH_4 \rightarrow HCl + \bullet CH_3$ reaction, DFT and HF/DFT-hybrids give a negative reaction barrier. The bond dissociation energy of O_2 calculated using DFT (GGA) is 15 kcal/mol in error. DFT is very useful for many systems in which it works, but no one is yet certain how to improve it systematically. Development of new functionals for transition states has also been hindered by the absence of reference data against which to fit the exchange-correlation functional.

Multi-reference and Post-Hartree-Fock Methods

Methods such as MCSCF, CC, CI, and CAS-MP2, are extremely accurate for activation barriers and can be systematically improved, but they are too expensive for all but the smallest systems. Hence there is a need to extend these advanced wavefunction methods to 100 atoms with large basis sets to handle problems of industrial relevance.

Locating Transition States

Locating transition states in complex systems is also a significant concern. In industrial applications, systems that contain multiple reaction paths, multiple steps, and multiple species are the norm, rather than the exception. For example, during the combustion process over 68 stationary points have already been identified on the reaction potential surface for the vinyl radical reacting with O_2, a four-heavy-atom problem (Westmoreland 2000). Hence there is a need to develop faster *ab initio* dynamics for larger systems, or new algorithms for searching a large reaction space.

Achieving Great Speed

Speed is critical for a wide variety of industrially important systems. For electronic structure calculations, linear scaling of cpu time with the number of electrons (N) (i.e., order(N)) is particularly important for strongly electron-correlated systems such as transition metals. Speed is also important for the simulation of many polymer properties such as folding, stress-strain, shear flow, and morphology after extrusion. Relaxation times need to be extended well past the currently feasible nanoseconds to 1 second and beyond. Only with greater speed and algorithm development can we hope to predict phenomena such as nonlinear mechanics and fracture mechanics of polymeric systems and their composites. Modeling long-term stability

with respect to slow processes such as oxidation, UV degradation, and physical aging for plastics also requires extensive simulation times, as failure is typically a rare-event process.

Porting Codes

Codes should be quickly ported to PCs and parallel systems to take full advantage of hardware advances.

Simulation of Interfaces

This represents another area of key challenges. Liquid-liquid interfaces are important for predicting phase diagrams, mesoscale structures, and the stability of micelles and emulsions. At the liquid-solid interface, accurate treatment of bonding, surface wetting, and transfer of electrons and atoms from one phase to another is essential for understanding corrosion, catalysis, thin-film deposition and surface passivation, for example.

Predicting Solution Properties of Complex Solvents and Mixtures

This is another difficult but important area in which progress is needed. For example, vapor-liquid equilibrium, particularly of complex mixtures, is a key issue for chemical manufacturing in general. Another example is the dispersion of large, irregular "molecular" species or aggregates such as asphaltenes or soot in oil. Even modeling the oil itself, which may contain over 5000 components ranging from long-chain hydrocarbons to polynuclear aromatics, is problematic and a major issue for the petroleum industry. Predicting the rheology of these dispersions and other complex systems such as polymeric solutions is still difficult. Polyelectrolytes introduce additional complexity, as the van der Waals and other dipole-dipole interactions fall off as $1/r^6$, whereas the electrostatic interaction energy is very long-range ($1/r$). Treating both interaction scales simultaneously is difficult because approximations that can be used for one scale result in a loss of necessary information for the other.

Greater Understanding

As the ability improves to do faster calculations on larger, more complex systems, the need for tools to enable interpretation, understanding, and communication of results increases as well. The ability to do a 1000-atom calculation in a day is not useful if the researcher cannot get value from the results. Hence visualization and analysis tools are essential to enable and communicate chemical understanding and learned "intuition." However, developing visualization tools and GUIs are not effective uses of a graduate student, post-doc, or industrial researcher's time. Hence a mechanism needs to be developed to make certain that this effort happens.

Improved tools to enable chemical understanding are also essential to move beyond expert-use-only, bringing many of these techniques to the bench chemists and engineers who can make routine use of them. For example, bench chemists are interested in quantum mechanical calculations to assess features such as bond strength, charge distribution, Fukui factors, and determining whether a reaction is charge or orbital controlled. However, the same person requires extensive analysis tools to visualize and interpret the calculated results in ways to which he or she is accustomed.

Manufacturing and Process Development

In addition to the needs of the chemical industry already mentioned, the manufacturing sector has additional requirements. The ultimate goal in process design is to be able to design efficient chemical processes from microscopic molecular behavior. During the next five years, it should be possible to make considerable progress by integrating what is already currently available. To date, there has been only minor integration of quantum mechanics with engineering applications through collaboration of personnel with separate tools within a company. To be effective, companies have expressed a need for greater integration of *ab initio* and molecular modeling with engineering modeling, both in commercially available applications and in academic departments and training of students. Experimental and calculated data need to be integrated as well.

Integrating Time and Length Scales

Over the next ten years, the chemical industry would like to see improved understanding of complex chemical processes through the integration and linking of simulation time and length scales. This is especially important for complex reaction systems that may require treating simultaneously multiple phases over multiple time scales, including exchange of material between phases. Today, tools exist at each level from the electronic to the continuum for modeling chemical reactions and processes, but they are not yet linked together or integrated.

Process Monitoring

Once a process has been designed, developed, and established, it is beneficial to develop the means to monitor the process so that it continuously produces material within specifications. As a continuous process wanders out of control, engineers need to know how to modify the process conditions to maintain the quality of the product. In many manufacturing processes, there is considerable lag time (hours+) between when a process change is made and when the impact on the product can be observed. Hence, there is considerable value in being able to simulate a process accurately enough to be able to predict *in real time* the impact of a change in processing conditions on the quality of the product.

Computing Facilities

Most companies have not yet opted for parallel-PC systems like the Beowulf cluster unless they can afford to have nearly one full-time person supporting it. Quantum chemistry does not parallelize well because atoms are not independent and separable from one another, requiring nearly continuous communication to calculate their interactions. Robert Harrison from Pacific Northwest National Labs has suggested that special-purpose computers such as those used in IBM's Blue Gene project will not carry the field of computational chemistry very far: "Computational chemistry as a whole is rich enough that only general-purpose computing is going to satisfy it" (C&EN 2000a).

SUMMARY

Through numerous successes, molecular modeling has become established as a valuable part of the R&D process in the chemical industry over the past five years. Success is very dependent on the abilities and training of the people doing the modeling, and their ability to work effectively with their experimental colleagues. It is not yet perceived as a necessity in all companies, but there is increasing demand for modeling results by researchers and for predictive models by formulators and engineers across the chemical industry. Prospects for growth are good.

REFERENCES

Chem. and Eng. News (C&EN). 2000. (March 13, p. 19).

Chem. and Eng. News (C&EN). 2000a. (April 24, p. 42).

Cummings, P.T., C.K. Hall, M. Neurock, A.Z. Panagiotopoulos, B.K. Peterson, and P.R. Westmoreland. 1998. *Report: NSF Workshop on Future Directions for Molecular Modeling and Simulation: Fundamentals and Applications*, Arlington, VA, November 3-4, 1997: National Science Foundation.

Douglas, J.M. 1988. *Conceptual Design of Chemical Processes*. New York: McGraw-Hill.

McCabe, C., D. Bedrov, G.D. Smith, and P.T. Cummings. 2001. Discriminating between correlations of experimental viscosity data using molecular simulation. *Ind. Eng. Chem. Res.* 40. 473-475.

Oleson, J.D. (ed.). 1996. *Technology Vision 2020: The U.S. Chemical Industry*, Washington, DC: American Chemical Society.

Siepmann, J.I., S. Karaborni, and B. Smit. 1993. Simulating the critical behaviour of complex fluids. *Nature*. 365 (6444):330-332.

Westmoreland, P.R. 2000. complexity of kinetics and product channels for the rection of vinyl reaction with oxygen. Symposium on Industrial Applications of Theoretical Chemistry. Paper PHYS-293. *220th National Meeting of ACS.* Washington, DC. August 20-24.

CHAPTER 6

APPLICATIONS: PHARMACEUTICALS AND LIFE SCIENCES

Peter A. Kollman

INTRODUCTION AND HISTORICAL OVERVIEW

The purpose of the pharmaceutical industry is to create new chemical entities that can combat disease and increase the quality of human life. This is a multi-billion dollar industry because in the late 20th and early 21st centuries, humans have high expectations about the quality of life. If feasible, they are willing to pay substantially for this. Since there is so much money to be made, the industry is very competitive, as is the scientific research in this industry. Given that a "blockbuster" drug can have sales in the billions of dollars for many years, the industry is willing to pay for science that gives them a competitive advantage, whether it is experimental or theoretical. Even though the time from beginning research to drug approval is of the order of 10 years, the pressure is to get the drugs in the pipeline as fast as possible to beat the competition. Thus, whatever methodologies are deemed useful to speed the process will be employed.

The cornerstone of research in the companies has been the synthetic organic chemist, because it is this scientist who must make the molecules that can be turned into drugs. Critical roles are played by experimentalists in the biological arena, e.g. molecular biologists to make proteins for in vitro testing and pharmacologists to create efficient and reliable tests for biological activity. The molecular modeling groups support the efforts to make the molecules that maximize the biological activity, and their roles will be described in more detail below.

Molecular modeling groups in the pharmaceutical industry began to appear in the late 1960s and early 1970s. These groups included those that employed Hansch's QSAR approach (Hansch and Fujita 1964), which was developed in the 1960s, the beginnings of molecular quantum mechanical calculations (Boyd 1998), and the earliest efforts in bringing molecular computer graphics to be used by the chemists (Gund et al. 1980). These groups really began to play a greater role and be more appreciated by their customers, the synthetic organic chemists, with the advent of interactive color computer graphics in the late 1970s (Langridge et al. 1981). The ability to visualize the three-dimensional properties of both active and inactive drug candidates, as well as (in an increasing number of cases) the target macromolecule to which the small molecules are bound, had the great potential to trigger new and more general ideas on what candidates should be synthesized to optimize binding activity. The active analog approach of Marshall (Sufrin, Dunn, and Marshall 1981) and co-workers represented the three-dimensional aspect of biological activity, just as the Hansch approach focused on physical properties like octanol-water partition coefficients.

Since that time, most pharmaceutical companies have had modeling groups, whose main purpose could be summarized thusly: Aid the experimental chemists and biologists to find a molecule that (a) binds tightly to its target, and (b) has excellent absorption, distribution, metabolism and excretion (ADME) properties.

THE APPROACHES TO MOLECULAR MODELING IN THE PHARMACEUTICAL AND BIOTECHNOLOGY INDUSTRIES

What are the scientific approaches that must be employed to find new chemical entities in the pharmaceutical industry? Since most drugs bind to their target non-covalently, through electrostatic and van der Waals/hydrophobic interactions, and these interactions dominate absorption, distribution and excretion properties, quantum mechanical electronic structure calculations play a secondary role in modeling in the pharmaceutical industry. To simulate the ADME properties, empirical models that have a large statistical component are used; to simulate the ligand-target interaction, molecular mechanics and other methods that focus on non-covalent interactions are adequate. Accurate representation of the ligand target interaction computationally requires (1) an accurate representation of non-covalent interactions due to ligand solvation, (2) non-covalent interactions involved in ligand conformational changes upon going from solvent to binding site, and (3) non-covalent interactions involved in ligand-target interactions. For reasons elaborated below, computationally demanding numerical calculations are rarely used in this process. This is because, in a nutshell, the number of possibilities for variation in the chemical nature of the ligand are so large that faster, approximate methods are more effective in guiding the synthetic chemist on how to proceed.

The methodologies used in the modeling depend on how much structural information is available for the target. There are basically three types of cases:

1. The structure of the target is known or can be reliably built using homology modeling. This situation is found in approximately 25% of modeling efforts and this number is growing. In this case, the modeler can use a molecular docking program such as DOCK (Sun et al. 1998) or FLEXX (Kramer et al. 1999) in order to find novel molecules that will bind to the target site. Both are fast enough to be able to screen entire company databases, hundreds of thousands of molecules, in a few days on a typical workstation. These methods use simple scoring methods because they must be very efficient, although with increasing computer power, modelers have begun to include solvation effects.

   The modeler then tests the best-scoring molecules and typically finds micromolar binders in this set, which one can then expand using two-dimensional (2D) and three-dimensional (3D) database searching in order to find related ligands and also try to crystallize and solve the structure bound to the target. Information from the three-dimensional structures of target-ligand and synthetic feasibility then is used to decide how to proceed to optimize the ligand. The power of this approach was first appreciated by Agouron and Vertex Pharmaceuticals, companies that were founded on the idea that this approach would be increasingly successful, given the increasing power of X-ray crystallography (and more recently NMR) to solve structures of macromolecules and their complexes. The fact that both companies came up with HIV protease inhibitors (Rao, Kim, and Murcko 1996; Kaldor et al. 1997) is impressive evidence that this concept works. Of course, this is a validation of structure-based drug design, not numerical modeling, in that computer graphics visualization and manipulation of the structure were clearly key components in this process. However, modeling was likely to have influenced the thinking of the synthetic chemists involved in this process.

2. The structure of the target is not known. This corresponds to approximately 50% of the modeling. In this case, a modeler will try to structurally superimpose the active molecules and inactive molecules and try to ascertain what are the key features that lead to activity. In a modern implementation of Marshall's active analog approach, Tripos, Inc. developed Comparative Molecular Field Analysis (CoMFA) (Tropsha and Cho 1998), in which the modeler analyzes electrostatic and steric fields around the active and inactive molecules and attempts to derive the key features that represent activity. Hydrophobicity or other properties can be considered in the analysis. The modeler can then search for 2D and 3D similarities with the actives from large databases of compounds and try to expand the list of actives.

3. The target is known to be a G-protein coupled receptor. A modeler uses the structure of bacteriorhodopsin (Luecke et al. 1999) or rhodopsin to build a model of the receptor, taking advantage of known site-specific mutations in the receptor and structure-activity data on the ligands to try to build a three-dimensional model of the target-ligand complex that is consistent with the data. The modeler can then do 2D and 3D searching for other ligands, trying to fit them into the model as actives or inactives and attempting continually to falsify the current model and improve it. This type of modeling corresponds to 25% of the cases, even though, in companies where no such modeling is done, the system

is treated as if no target structure is known (category 2 above). The bacteriorhodopsin structure is a 7 transmembrane helix protein and the G-protein coupled receptors are known to be of this class, but it is not clear how reliably one can build the new sequences into this structure. Given that one cannot get a three-dimensional structure of the target-ligand complex, one is "flying a bit blind" in using such an approach. Nonetheless, with enough structure/activity data on ligand and receptor residues, such models can be useful to the organic chemist thinking about new ligands to synthesize.

STRUCTURE AND DUTIES OF THE MODELING GROUPS IN PHARMACEUTICAL AND BIOTECH COMPANIES

In the modeling groups we visited, the number of group members was typically 3-10% of the number of synthetic organic chemists in the company. Their main responsibility was to provide computational support for the drug design projects. In this guise, the modeler was part of a project team, consisting of synthetic chemists and biologists also, that met regularly to share results and discuss strategies. Each modeler was expected to provide support for 1-2 major projects, although they might have some other duties as well. In the larger pharmaceutical companies, part of the modeling group had responsibilities for software development because it was felt that special homegrown software was important to give the company a competitive advantage.

Typical training for the members of the modeling groups included graduate degrees in quantum chemistry, computational chemistry, macromolecular crystallography and biophysics. In addition, it is clear that for those working closely as part of the project teams, an excellent understanding of organic chemistry is critical, both for the modeling and for communicating with the synthetic organic chemists. For the members of the modeling group whose duty involves more software development, more computational or computer science background is obviously likely to be helpful.

The nature of the modeler-synthetic chemist interactions varied widely, not only between but also inside a company. The synthetic chemists' receptiveness to modeling tended to be greater for recent PhDs who have some knowledge of modeling from their own education. Of course, when modeling has led to some successes, as in some companies we saw, the active participation and interest of the organic chemist in using the fruits of modeling is strongly encouraged by the organic chemists who are the group leaders.

Over the years and between companies, one has seen two extreme attitudes toward modeling: (a) let the modelers do the modeling and the synthetic chemist make more molecules, or (b) every synthetic chemist should spend time in front of the graphics terminal. Of course these two extremes are not so different. The modeling group can work on software or more demanding calculations, and the synthetic chemists can use the fruits of modern molecular visualization tools, which do not require so much training in computational approaches.

There are also two organizational models we observed, one in which the modelers were farmed out to the therapeutic areas and reported to the leaders of those groups, the other where the modeling group reported to the leader of modeling or physical chemistry. Both models can work, given the correct attitude, environment and expectation, but it is clear that even if the reporting is to therapeutic areas, locating the modelers together leads to much useful synergy in techniques and ideas brought to the therapeutic projects.

HARDWARE AND SOFTWARE IN MODELING GROUPS IN PHARMACEUTICAL COMPANIES

The types of software required for modeling efforts can be listed in approximate order of importance to the companies, first giving examples of commercial and academic codes and then describing the commercial companies whose software is most used:

- The single most ubiquitous use for modeling software is in computer graphics, where Insight (Acclrys / MSI) and Sybyl (Tripos) are the most frequently used. The academic code MIDAS is a typical academic code in this area.

- Secondly, software tools for large-scale searching for 2D/3D matches (distributed by commercial companies) are much used, as are various statistical approaches, 3DQSAR, such as COMFA (Tropsha and Cho 1998).
- Third, there are molecular docking codes such as DOCK (Yun et al. 1998), developed at UCSF, and FLEXX (Kramer et al. 1999), developed in an academic group in Germany and distributed by Tripos Inc.
- Fourth, there is software for combinatorial chemistry library generation and analysis, distributed by commercial vendors and also with significant in-house development at the companies.
- Fifth are molecular mechanics/dynamics, such as Discover (Ewig, Thatcher, and Hagler 1999) and CHARMM (Brooks et al. 1983), distributed by Accylrys / MSI, and AMBER (Pearlman et al. 1995) and GROMOS (van Gunsteren, Daura, and Mark 1998), academic codes from UCSF and ETH Zürich, respectively.
- Finally, there are quantum chemistry codes, such as Gaussian XX (Frisch et al. 1998), MOPAC (Stewart 1990), Spartan (Hehre et al. 1991), and HyperChem (Froimowitz 1993).

Companies spend between $200,000 and $1,000,000 per year on software and have the equivalent of 1-10 SGI 10000 equivalents per member of the modeling group. There is an appreciation for the price/performance for PC clusters, but no large-scale embracing of this approach at this time.

Different companies devote varying amounts of effort to the development of their own computer software. On one hand, it is realized that such development can give a company a competitive advantage. There is an effort at Taisho in Japan to develop specialized hardware for the same reason. On the other hand, such in-house software and specialized hardware is expensive to develop and maintain. Even given the large amounts of money that can result from a big success, it can be challenging to justify the expense of such a project if the company does not have a history of useful modeling successes and farsighted leaders.

Commercial software in the pharmaceuticals area is dominated by Accylrys / MSI and Tripos. Molecular Design and Daylight are extensively used in chemical databases and data management. There is an effort in Japan to develop software, partly in order not to become too dependent on American companies and their pricing.

RELATIONSHIP OF MODELING TO OTHER TECHNOLOGIES

In most companies, there are close connections between the modeling and X-ray crystallography groups, in companies that have these groups. Given the essential role of structures in modeling and the fact that much more precise modeling can be done in this case, many companies have extensive investments in X-ray crystallography and many in NMR spectroscopy as well. It is obvious that there can be great synergy between the crystallography and modeling groups, and at Takeda, they are combined in one group.

Combinatorial chemistry (Kirkpatrick, Watson, and Ulhaq 1999) is having a major impact on drug discovery, since one can make more molecules by this approach than one could imagine prior to the development of this technique. At first, this technology was thought to be the revenge of the synthetic chemists, since they could make everything and not have to rely on modeling, but now it is clear that one can't make everything, so that one should be very clever about what to make. In many modeling groups, extensive involvement in combinatorial library design and planning is taking place. For example, the Combichem Company (now a part of Bristol-Myers Squibb through its acquisition of DuPont Pharmaceuticals) had separate groups for software development, combinatorial chemistry and modeling applications.

The era of genomics (Teichman, Chothia, and Gerstein 1999) is upon us, and SmithKline Beecham (now part of Glaxo SmithKline) was among the earliest companies to embrace genomics. The number of potential drug targets is enormous and can come both from the full sequencing of pathogen genomes and from the human genome. In many companies, there are separate bioinformatics groups in order to handle the tremendous amount of potential information and possibilities inherent in so many new targets for potential drug discovery.

The combination of combinatorial chemistry and genomics is likely to continue to increase the number of potential small molecules and targets that must be considered in drug discovery. This is going to make it difficult to use the increased computer power inherent in Moore's law to significantly increase the accuracy of typical calculations involved in the drug discovery process, since the calculations will have to be carried out on so many more ligands and targets. Nonetheless, the opportunity for cleverness in wading through this thicket of possibilities is enormous and certain to be a great challenge for the modeling groups in the industry.

OVERVIEW OF ROLE, SUCCESSES AND FAILURES OF MODELING GROUPS

Already now, and certain to be increasingly so in the future, the key role of the modeling groups is to identify which molecules should be made and, equally importantly, *not* made in the process of finding a lead compound. Further testing can then evaluate whether it should be turned into a Safety Assessment Candidate (SAC), which is a key step to turning it into a drug.

Although accurate calculations are desirable in this process, it appears more important to be able to use calculations which, albeit less accurate, can consider many more possibilities. Nonetheless, there are opportunities for clever use of increasingly accurate calculations after simpler ones have been used to filter the possibilities to a number manageable by the more demanding calculations. It is interesting that the most extensive use of the more computationally demanding calculations is found in the structure-based design companies such as Agouron and Vertex Pharmaceuticals (Ajay and Murcko 1995; Pearlman 1999; Reddy and Erion 1998; Retio, Bouzida, and Verkhivker 1999). Thus, the modeling efforts improve their odds of catalyzing the discovery of an SAC depending on how clever the modelers are in covering the many possibilities in "chemical space" and presenting this information to the synthetic organic chemist in a useful way. This may involve writing unique software, using existing software, or a combination of the two.

One of the points noted in this WTEC panel's visits was that the modeling group can play a unique role as technical "gate-openers" in larger pharmaceutical companies. Its members cut across project teams, while the typical synthetic chemist and biologist are members of only one team at any one time. Ideally, in this role the modeler can transmit useful ideas from one project team to another.

We saw a number of examples of successes of modeling during our visits. For example, at Mitsubishi, a combination of in-house software and DOCK led to a much greater number of hits from a database than would have been found by random screening of this database. Similar results from DOCK have been found at a number of other companies (Kuntz 1999). The modeling group also played an educational role. For example, at SmithKline Beecham in England the concept of "modeling chemists" was used in a minicourse put on to educate the synthetic chemists on the available capabilities of modeling. At DuPont Pharmaceuticals, we met a leader of an organic chemistry team, who, because this approach had led to some exciting molecules before, strongly encouraged his synthetic chemists to spend time in front of the graphics terminal, thinking about what molecules to make next.

On the other hand, we heard expressions of disappointment that the modeling hadn't reached the level of usefulness of, for example, molecular dynamics refinement of crystallographic structures in the software XPLOR (Brunger, Kuriyan, and Karplus 1987), which has transformed refinement of macromolecular structures. Of course, XPLOR uses experimental X-ray reflections to guide its refinement also, while pure modeling must come to grips with the myriad of chemical and conformational spaces without direct guidance from experiment other than a few macromolecular structures.

NEEDS FOR THE FUTURE

The most common need expressed during the WTEC panel's visits was for improvements in DOCKing and evaluating the DOCK leads with a scoring function. If the number of hits could be significantly increased and lead more rapidly to more tight binding hits, this would have a major impact on the drug discovery process. Of course, this improvement would only apply to modeling in the first category, where the target structure is known, but the number of such cases is bound to increase over the years.

Of course, the ability to accurately determine the three-dimensional structures of G-protein coupled receptors would have an enormous impact on modeling in pharmaceutical design, but this is not something that can be "forced"; membrane protein crystallization has to date been very difficult. However, there has been a recent solution of the X-ray crystal structure of bovine rhodopsin that bodes well for future developments in the structural biology of G-protein coupled receptors (Kalczewski et al. 2000).

The second key element in drug design, ADME properties, is essential to represent, but there is not a good theoretical way to do so. Thus there is a need for empirical models. For example, Hoffmann-LaRoche in Basel is putting a large effort into empirically determining these properties for a very large number of their compounds. It is clear that one way to improve the timeline for drug discovery is to be considering ADME properties from the beginning, rather than to first optimize binding to the target and then to consider ADME. Improved models and approaches to ADME properties would likely have a very large impact on molecular modeling in the drug discovery process.

Other elements of basic science which it was felt would have an impact on modeling in the pharmaceutical industry are improved force fields for molecular mechanics, improved solvation models, more accurate coupling of quantum and molecular mechanical models, and improved prediction of three-dimensional proteins models from sequences.

One of the trends in pharmaceutical modeling is to consider greater and greater numbers of possibilities, both in terms of combinatorial chemistry and in terms of genomics. Thus, at one company we heard the desirability of improved training for students in statistical concepts, clearly a necessity when dealing with so many possibilities. A thorough enough grounding in needed statistics is not often part of the training in computational chemistry. In a related comment, one site visit host noted that the discipline of computational chemistry should be broadened beyond the previous confines of quantum chemistry and molecular mechanical simulation approaches to include bio- and chemical informatics (and presumably statistics also). Thus, although it is very hard for one person to have all this expertise, it is clear that molecular modeling groups in the pharmaceutical industry must have a broad range of expertise to be most effective.

INTERNATIONAL COMPETITIVENESS IN THE PHARMACEUTICAL INDUSTRY

It is hard to separate U.S. from European companies because of the multinational nature of many of the companies we visited. Japanese pharmaceutical companies tend to be smaller, both in terms of numbers of synthetic chemists and in terms of numbers of modelers, but the percentage of modelers is not necessarily that much smaller relative to the synthetic chemists. The software used is predominantly from U.S. commercial vendors, although most companies have a reasonable amount of in-house software. In the related sciences, i.e., those that relate to pharmaceuticals, the United States seems to have a higher level of activity than Europe, which has a higher level of activity than Japan.

Even though, of the companies discussed in this paragraph, the panel visited only Pharmacopeia, the author of this chapter has had some contacts or knows people at the other companies, so some information has been gleaned informally. One of the things that stands out is the fertile environment in the United States in recent years for startup companies in niche technologies, such as structure-based design. Agouron and Vertex have, by most measures, been quite successful both in their own researches and by forming alliances with other, more traditional pharmaceutical companies. It is too early to tell for combinatorial chemistry companies, but Arqule, Pharmacopeia and Combichem (now a part of Bristol-Myers Squibb) have also been able to make inroads in the pharmaceutical industry. Thus, one of the characteristics of the United States, less so in Europe and least of all in Japan, is the agility with which companies in specialized new technologies in drug design spring up. Agouron and Vertex have large and impressive modeling groups that support their efforts in structure-based drug design; Pharmacopeia has two modeling groups, one to support its own in-house design and one to test out new technologies to complement its other half, Accylrys / MSI; Combichem had two modeling groups, one developing software and the other applying it, to complement their experimental combinatorial chemistry group.

Modeling in the pharmaceutical industry is very vital and exciting. It is changing its nature to include bioinformatic and cheminformatic components and the revolution in combinatorial chemistry and genomics.

This will be a great challenge for the modeling groups—to find efficient ways of considering many more possible organic molecules and targets, while providing accurate estimates for binding and ADME properties of the most promising leads.

The panel did not visit any purely agrochemical companies or discuss toxicology with anyone in any detail. DuPont was the only site where agricultural chemicals were discussed, and a subsequent paper gives more details about DuPont's work in this field (Kleier 2001). A site report in Appendix B on Bayer notes its activity in the field. However, the problems in agrochemicals are analogous to those in pharmaceuticals, although the requirement for very inexpensive molecules puts more constraints on design. The issue of toxicology is closely analogous to the issue of ADME properties noted above. Because of the interest in agrochemicals, the era of genomics is likely to enable selective targeting of new genes, e.g., in pesticides, and also more structure determinations of pesticide-specific target proteins.

REFERENCES

Ajay and M.A. Murcko. 1995. *J. Med. Chem.*, 38, 4953.

Boyd, D. 1998. Drug design. In *Encyclopedia of Computational Chemistry*, P.v.R. Schleyer (ed.), J. Wiley, Chichester, vol. 1, p. 795.

Brooks, B.R., R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. 1983. *J. Comp. Chem.*, 4, 187.

Brunger, A.T., J. Kuriyan and M. Karplus. 1987. *Science*, 235, 458.

Ewig, C.S., T.S. Thatcher and A.T. Hagler. 1999. *J. Phys. Chem. B*, 103, 6998.

Frisch, M.J., et al. 1998. Gaussian Inc. Pittsburgh, PA. (http://www.gaussian.com).

Froimowitz, M. 1993. *Biotechniques*, 14, 1010. and HyperChem™. Hypercube, Inc. 1115 NW 4th Street, Gainsville, FL 32601. (http://www.hyper.com).

Gund, P., J.D. Andose, J.B. Rhodes, and G.M. Smith. 1980. *Science*, 208, 1425.

Hansch, C., and T. Fujita. 1964. *J. Amer. Chem. Soc.*, 84, 1616.

Hehre, W. et al. 1991. Wavefunction Inc. Irvine, CA, 92612 . (http://www.wavefun.com; founded 1991).

Kalczewski, K. et al. 2000. *Science*, 289, 739-745.

Kaldor, S.W., V.J. Kalish, J.F. Davies, B.V. Shetty, et al. 1997. *J. Med. Chem.*, 40, 3979.

Kirkpatrick, L., S. Watson and S. Ulhaq. 1999. *Combinatorial Chemistry and High Throughput Screening*, 2, 211.

Kleier, D.A. 2001. Molecular modeling and simulation of crop-protection chemicals. In *Foundations of Molecular Modeling and Simulations*. P.T. Cummings, P.R. Westmoreland, B. Carnahan (eds.). AIChE Symp. Series No. 325, vol. 97. AichE. New York. p. 9.

Kramer, B., G. Metz, M. Rarey and T. Lengauer. 1999. *Med. Chem. Res.*, 9, 463.

Kuntz, I.D. 1999. Personal communication of results reported at the DOCK users group meeting, UCSF.

Langridge, R., T.E. Ferrin, I.D. Kuntz, and M.L. Connolly. 1981. *Science*, 211, 661.

Luecke, H., B. Schobert, H.J. Richter, J.P. Cartailler, and J.K. Lanyi. 1999. *J. Mol. Biol.*, 291, 899.

Pearlman, D.A. 1999. *J. Med. Chem.*, 42, 4313.

Pearlman, D.A., et al. 1995. *Comput. Phys. Comm.*, 91, 1.

Rao, B.G., E.E. Kim and M.A. Murcko. 1996. *J. Comp. Aided Mol. Des.* 10, 23.

Reddy, M.R., and M.D. Erion. 1998. *J. Enzyme Inhibition*, 14,1.

Retjo, P.A., D. Bouzida and G.M. Verkhivker. 1999. *Theor. Chem. Acc.*, 101,138.

Stewart, J.J.P. 1990. *J. Comp. Chem.*, 10, 209.

Sufrin, J.R., D.A. Dunn and G.R. Marshall. 1981. *Mol. Pharmacol.*, 19, 307.

Sun, Y., T.J.A. Ewing, A.G. Skillman and I.D. Kuntz. 1998. *J. Comp. Aided Mol. Des.*, 12, 597.

Teichman, S.A., C. Chothia and M. Gerstein. 1999. *Curr. Opin. Struct. Biol.*, 9, 390.

Tropsha and S.J. Cho. 1998. *Perspect. Drug Discovery Design*, 12, 57 (see also http://www.tripos.com).

van Gunsteren, W., X. Daura and A.E. Mark. 1998. GROMOS force field. In *Encyclopedia of Computational Chemistry*, P.v.R. Schleyer (ed.), J. Wiley, Chichester, v. 2, p. 1211.

CHAPTER 7

APPLICATIONS: PHYSICAL AND ELECTRONIC MATERIALS

Priya Vashishta

INTRODUCTION

Research and development of materials have thus far focused on controlling structures at diverse length scales—atoms, defects, fibers, interfaces, grains, pores. Because of the inherent complexity of such multiscale materials phenomena, computer modeling is expected to play an important role in the design of materials such as metals, semiconductors, ceramics, and glasses (Pechenik et al. 1999). Recent years have witnessed rapid progress in large-scale atomistic simulations, highly efficient algorithms for massively parallel machines, and immersive and interactive virtual environments for analyzing and controlling simulations in real time (Abraham 1997; Germann and Lomdahl 1999; Kalia et al. 2000; Vashishta et al. 1999). As a result of these advances, simulation efforts are being directed toward reliably predicting properties of materials in advance of fabrication. Thus, materials simulations are capable of complementing and guiding experimental search for novel materials.

A comparative assessment of the science of materials modeling and its industrial applications in the United States, Europe, and Japan in the areas of simulation methods and applications—metals, semiconductors, ceramics, glasses, interfaces, carbon systems, and supercritical water—is discussed in the section "Comparison of science and industrial application of science in the United States, Europe and Japan" of this chapter.

Site reports in Appendix B also provide further details. Relevant site reports from the United States are as follows:

- Lucent Technologies (www.lucent.com)
- Motorola (www.motorola.com)
- Sandia National Laboratories (www.sandia.gov)
- Air Products (www.airproducts.com)
- National Institute of Standards and Technology (www.nist.gov)
- University of Minnesota Supercomputing Center (www.msi.umn.edu)
- LSU Concurrent Computing Laboratory for Molecular Simulations (www.cclms.lsu.edu)

Research at IBM Yorktown Heights and IBM San Jose (www.ibm.com), and Texas Instruments (www.ti.com) is quite similar to that at Lucent and Motorola. Site reports most relevant in Europe are as follows:

- Daresbury Laboratories, UK (www.dl.ac.uk)
- IBM-Zürich (www.zurich.ibm.com)
- TU-Denmark (www.dtu.dk)

Complementary to the work at Daresbury Laboratories are the following more academic activities:

- Cambridge University (www.cam.ac.uk) by M. Payne, N. Handy, J.-P. Hansen, D. Wells, M. Sprik, and P. D. Bristowe
- Oxford University (www.ox.ac.uk) by P. A. Madden, J. M. Yeomans, K. Refson, and D. Pettifor
- Max-Planck-Institut für Festkörperforschung, Stuttgart (www.mpi-stuttgart.mpg.de), by M. Parrinello and O. K. Andersen
- Johannes Gutenberg University of Mainz (R. Schilling, K. Binder, www.cond-mat.physik.uni-mainz.de/) and Max Planck Institute for Polymer Research (K. Kremer, www.mpip-mainz.mpg.de/)
- Delft University, The Netherlands (www.tudelft.nl) by S. de Leeuw

Japanese site reports most relevant for this chapter are the following:

- Fujitsu (www.fujitsu.co.jp)
- Institute of Solid State Physics, Tokyo (www.issp.u-tokyo.ac.jp)
- JRCAT, Tsukuba (www.jrcat.or.jp)
- NEC Research Laboratory, Tsukuba (www.nec.co.jp)
- Toshiba (www.toshiba.co.jp)

Activities complementary to the work at JRCAT may be found at the following sites:

- Niigata University (www.niigata-u.ac.jp),
- Okayama University (www.okayama-u.ac.jp)
- Hiroshima University (www.hiroshima-u.ac.jp)
- Yamaguchi University (www.yamaguchi-u.ac.jp)

METHODS

Molecular Dynamics

Molecular-dynamics (MD) simulations have played a key role in our understanding of material properties and processes. In the MD approach (see Chapter 3), one obtains the phase-space trajectories of the system (positions and velocities of all the atoms at all time) from the numerical solution of Newton's equations (Allen and Tildesley 1987; Rapaport 1995; Smit and Fraankel 1996). This allows one to study how atomistic processes determine macroscopic materials properties. Recent advances in scalable, space-time multiresolution algorithms coupled with access to massively parallel computing resources have enabled large, multimillion-atom (1-million to 1-billion atoms) MD simulations of complex materials.

Interatomic Potentials

Accurate atomic force laws are essential for realistic molecular and materials simulations of real materials and processes. Mathematically, a force law is encoded in the interatomic potential energy. In the early years of simulations, the most frequently used interatomic interaction was the Lennard-Jones potential. For ionic materials, the interaction potentials included long-range Coulomb interaction. Most of these interatomic interactions fell in the category of simple two-body potentials. For silicon, which is a covalent material, the most successful potential function is due to Stillinger and Weber (1985).

In past years, reliable interatomic potentials have been developed for a number of materials, including ceramics such as silica (SiO_2), silicon nitride (Si_3N_4), and silicon carbide (SiC), as well as semiconductors such as gallium arsenide (GaAs), aluminum arsenide (AlAs), and indium arsenide (InAs) (Vashishta et al. 1996). Interatomic potential energy for these materials consists of two- and three-body terms. The two-body potential energy is a sum over contributions from $N(N+1)/2$ atomic pairs, (i, j). The contribution from each pair depends only on their relative distance, $|\boldsymbol{r}_{ij}|$. Physically, the two-body terms are steric repulsion and electrostatic interaction due to charge transfer between atoms, and charge-dipole and dipole-dipole interactions that take into account the large electronic polarizability of negative ions. The three-body potential energy consists of contributions from atomic triples (i, j, k), and takes into account covalent effects through bending and stretching of atomic bonds, $\boldsymbol{r}_{ij}$ and $\boldsymbol{r}_{ik}$.

These analytic force models are often derived from first-principles calculations. The database to be used to parameterize the interatomic potentials can be generated from first-principles total energy calculations on key structures, surface and point defects, and adhesive energies between various interfaces.

Recent schemes also include environment-dependent (rather than fixed) partial charges to model charge-transfer and the resulting long-range Coulomb interaction (Campbell et al. 1999; Ogata et al. 1999; Rappé and Goddard 1991; Rick et al. 1994; Streitz and Mintmire 1994). In a semi-empirical approach, partial charges associated with each atom are determined from an environment-dependent electronegativity and self-Coulomb repulsion, each calculated from the atomic Slater-type orbitals. For metals, the embedded-atom method (EAM) provides a reasonable description of interatomic forces (Foiles et al. 1986), while for group IV elements, Tersoff-type bond-order potentials capture most aspects of bonding and atomic hybridization (Bazant et al. 1997; Brenner 2000; Pettifor and Oleinik 2000; Tersoff 1986).

A reactive empirical bond-order potential, which is based on the second-moment approximation to the local electronic density of states, models the interatomic energy as a sum of repulsive and attractive pair potentials (Brenner 2000). The attractive pair terms, which model bonding from the valence electrons, are modulated by empirical bond order functions whose values depend on coordination, bond angles, radical character of the bond, and an approximation to conjugation effects arising from adjacent unsaturated atoms.

Quantum Mechanics

Empirical interatomic potentials used in MD simulations fail to describe chemical processes. Instead, interatomic interaction in reactive regions needs to be calculated by a quantum-mechanical (QM) method that can describe breaking and formation of bonds. An atom consists of a nucleus and surrounding electrons, and QM schemes treat electronic degrees-of-freedom explicitly, thereby describing the wave-mechanical nature of electrons.

One of the simplest QM schemes is based on the tight-binding method (Wang and Ho 1996). The TB method does not involve electronic wave functions explicitly, but rather it solves an eigenvalue problem for the matrix that represents interference between electronic orbitals. The spectrum of the eigenvalues gives the information on electronic density of states. In the TB scheme, electronic contribution to interatomic forces is derived through the Helmann-Feynman theorem, which states that only partial derivatives of the matrix elements with respect to $\vec{r}^{\,N}$ contribute to forces.

A more accurate but computationally intensive QM scheme deals explicitly with electronic wave functions, $\psi^{N_{wf}}(\mathbf{r}) = \{\psi_1(\mathbf{r}), \psi_2(\mathbf{r}), ..., \psi_{N_{wf}}(\mathbf{r})\}$ (N_{wf} is the number of independent wave functions, or electronic bands, in the QM calculation), and their mutual interaction in the framework of the density functional theory (DFT) (Hohenberg and Kohn 1964; Kohn and Vashishta 1983; Payne et al. 1992) and electron-ion interaction using pseudopotentials (Troullier and Martins 1991). The DFT, for the development of which Walter Kohn received a 1998 Nobel Prize in Chemistry, reduces the exponentially complex quantum N-body problem to a self-consistent eigenvalue problem that can be solved with $O(N_{wf}^3)$ operations. In the DFT scheme, not only are accurate interatomic forces obtained from the Helmann-Feynman theorem (Car and Parrinello 1985), but also electronic information such as charge distribution can be calculated.

Empirical interatomic potentials used in MD simulations fail to describe chemical processes. Capturing quantum mechanics and dynamics simultaneously is another approach. For example, a QM cluster may be embedded in an MD simulation (Singh and Kollman 1986; Field et al. 1990; Dapprich et al. 1999; Eichinger et al. 1999; Ogata et al. 2001).

Scalable Algorithms

Efficient algorithms are key to extending the scope of simulations to larger spatial and temporal scales that are otherwise impossible to simulate. These algorithms often utilize multiresolutions in both space and time (Nakano et al. 1999).

The most computationally intensive problem in an MD simulation is the computation of the electrostatic energy for N charged atoms. Direct evaluation of all the atomic-pair contributions requires $O(N^2)$ operations. In 1987, an $O(N)$ algorithm called the "fast multipole method" (FMM) was discovered (Greengard and Rokhlin 1987). The FMM groups distant atoms together and treats them collectively (Toukmaji and Board 1996). Hierarchical grouping is facilitated by recursively dividing the physical system into smaller cells, thereby generating a tree structure. By computing both expansions recursively for the hierarchy of cells, the electrostatic energy is computed with $O(N)$ operations. The FMM also has well-defined error bounds.

The discrete time step, Δt, in MD simulations must be chosen sufficiently small so that the fastest characteristic oscillations of the simulated system are accurately represented. However, many important physical processes are slow and are characterized by time scales that are many orders-of-magnitude larger than Δt (Schlick et al. 1997; Deuflhard et al. 1999). Molecular dynamics simulations of such "stiff" systems require many iteration steps, and this severely restricts the applicability of the simulation. An approach called the multiple time-scale (MTS) method uses different Δt for different force components to reduce the number of force evaluations (Martyna et al. 1994; Nakano et al. 1994; Skeel et al. 1997). To further speed up simulations, it is also possible to use a hierarchy of dynamics including rigid-body (or fuzzy-body) motion of atomic clusters (Nakano 1997).

Parallel computing technology has extended the scope of computer simulations in terms of simulated system size. Parallel computing requires decomposing the computation into subtasks and mapping them to multiple processors. For MD simulations, the divide-and-conquer strategy based on spatial decomposition is commonly used (Rapaport 1995).

Many MD simulations are characterized by irregular atomic distribution. Simulation of dynamic fracture is a typical example. One practical problem in simulating such irregular systems on parallel computers is that of load imbalance (Fox et al. 1994). Because of the irregular distribution of atoms, this uniform spatial decomposition results in unequal partition of workloads among processors. As a result the parallel efficiency is degraded significantly. This load-imbalance problem can be solved by partitioning the system not in the physical Euclidean space but in a computational space, which is related to the physical space by a curvilinear coordinate transformation (Nakano 1999). (The computational space shrinks where the workload density is high and expands where the density is low, so that the workload is uniformly distributed.) The optimal coordinate system is determined to minimize the load-imbalance and communication costs. Optimization is carried out using a simulated annealing method.

A serious technological gap exists between the growth in processor power and that of input/output (I/O) speed. The I/O (including data transfer to remote archival storage devices) has thus become the bottleneck in large-scale multimillion atom MD simulations. The I/O problem has been addressed using a scalable data-compression scheme (Omeltchenko et al. 2000). It uses octree indexing and sorts atoms accordingly on the resulting space-filling curve. By storing differences between successive atomic coordinates, the I/O requirement with the same error tolerance level reduces from $O(N\log N)$ to $O(N)$. This, together with variable-length encoding to handle exceptional values, reduces the I/O size by an order of magnitude with a user-controlled error bound.

APPLICATIONS

Recent developments in MD simulations have enabled researchers to study a number of materials properties: structure, vibrational and optical properties, thermal properties (thermal conductivity), mechanical properties (elastic moduli, toughness, hardness), and tribological properties. Materials processes that can be studied with MD simulations include growth (molecular-beam and chemical-vapor depositions), sintering, oxidation, fracture, indentation, pressure- and temperature-induced structural transformations, and impact damage. Quantum mechanical simulations further allow the study of electronic, magnetic, and optical properties from first principles.

The field of molecular and materials modeling is extremely broad. To elucidate the scope of materials modeling research and its industrial applications, the author has chosen some of the recent non-industrial simulations that highlight the relevance of atomistic modeling of materials to industrial interests, complementing the examples in the site reports. The simulations are of metals, ceramics and glasses, interfaces, and carbon nanostructures.

Metals

Novel multiscale materials modeling tools have recently been applied to unravel the fundamental physics behind the scientifically and technologically interesting phenomenon of dislocation dynamics in metals (de la Rubia et al. 2000). The work has been carried out by scientists at Lawrence Livermore National Laboratory, led by Dr. Tomas Diaz de la Rubia of the Materials Science and Technology Division.

Materials under stress often exhibit inhomogeneous deformation patterns resulting from localization of the plastic flow. This phenomenon, which ultimately leads to premature failure and occurs in many classes of materials under a wide variety of deformation conditions, has been known and studied for over 40 years. The following are examples:

- Line defects called dislocations are known to form regular patterns (cell walls) during the early stages of deformation of metals.
- Shear bands appear during high strain rate deformation and forming of materials.
- Localization and liquefaction is known to occur in geological materials.
- Recently it has been shown that localization of plastic flow also occurs in deformed bulk amorphous metals.

Despite the broad importance of the problem, no clear description of the microscopic mechanisms that give rise to the inhomogeneous macroscopic behavior exists. Previous theories have been based on continuum level phenomenological constitutive laws and have not provided a connection between the microscopic degrees of freedom and the macroscopic behavior.

De la Rubia and coworkers employed three-dimensional dislocation dynamics (DD) to investigate plastic flow localization during deformation of irradiated materials. The DD model is a powerful tool for investigating the mechanical response of materials on a mesoscopic scale. In a DD simulation, the plastic deformation of a single crystal is obtained by explicit accounting of the dislocation evolution history, i.e., their motion and structure. The motion and interaction of an ensemble of dislocations in a three-dimensional crystal is marched in time. A typical simulation domain is a box of side 10-30 μm. Dislocations are discretized into straight-line segments, and the stress field produced by each segment is calculated. The Peach-Koehler force F acting on a dislocation segment inside the computational cell is calculated from the stress fields due to immediate neighboring segments, all other dislocations, all defect clusters, and the applied stress. The result is used to advance the dislocation segment. The problem is computationally complex and intense owing to the long-range $1/r$ stress field of dislocations and to the extreme complexity of the physics associated with collective behavior of large ensembles of these defects.

Currently, these DD simulations are being carried out at Lawrence Livermore National Laboratory with the *micro3D* code developed originally by Zbib and coworkers at Washington State University. *Micro3D* runs on the ASCI Blue supercomputer. It is capable of modeling the behavior of thousands of dislocation lines

simultaneously on crystals with microscopic dimensions on the order of tens of microns. At present, the maximum total plastic deformation that they have been able to achieve on a single crystal of molybdenum is 0.1%. While this is well beyond the elastic limit of the material, much progress in algorithm development is still required to take advantage of the upcoming generation of ASCI-class supercomputers with thousands of 16-processor nodes. Efficient use of these new massively parallel machines should enable achievement of plastic strains in excess of 10%, making the simulations more realistic, predictive, and easier to compare to experiments.

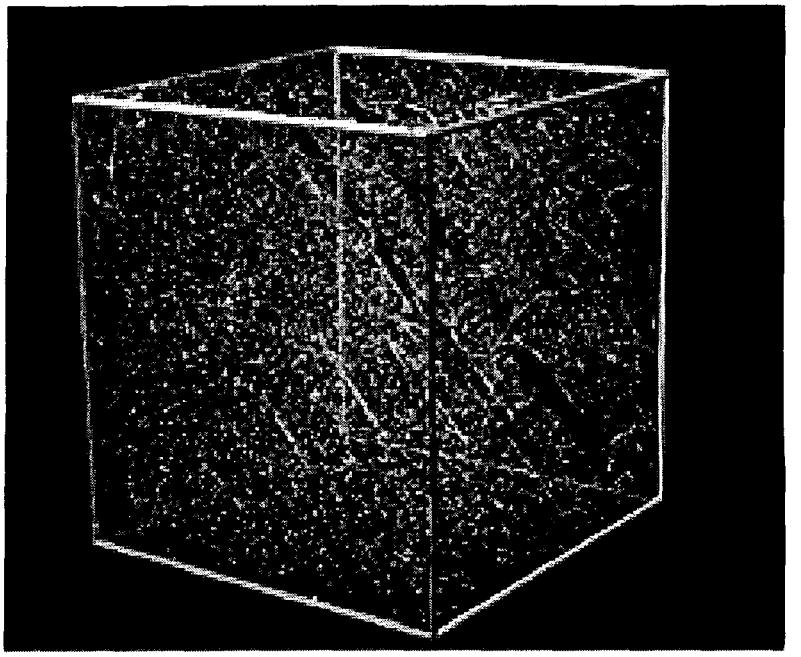


Fig. 7.1. Results of a DD simulation of plastic flow localization in irradiated copper. This was the first-ever simulation that demonstrated the physical mechanisms underlying this process (de la Rubia et al. 2000).

Ceramics and Glasses

Advanced structural ceramics are highly desirable materials for applications in extreme operating conditions. Light weight, elevated melting temperatures, high strengths, and wear and corrosion resistance make them very attractive for high-temperature and high-stress applications. The only serious drawback of ceramics is that they are brittle at low to moderately high temperatures.

A number of methods have been proposed for processing/synthesizing ceramic materials in new forms. For example, nanophase ceramics consisting of sintered/consolidated nanometer-size particles and fiber-reinforced ceramics are found to be fracture-tough; however, these materials have low creep resistance (Siegel 1996). For high-temperature applications, an ideal ceramic material should have high fracture-toughness and high creep resistance. The important point to emphasize here is that materials that are chemically nearly identical can have very different mechanical behavior depending on the method of processing. It is in this area that we expect large-scale simulations to play an important role.

The ceramics industry is most advanced in Japan. Europe and the United States also have some strong ceramics companies, but the breadth is not so great as in Japan. In Japan ceramics research is located in ceramic industry and government research institutes (Nagoya). In the United States, ceramics research programs are located at DOD and DOE laboratories. Modeling of ceramic materials is in its early stages.

Until recently, there was not enough computer power to carry out simulations of processing. Large-scale computer simulations will for the first time permit the "What if?" type questions about mechanical properties when the processing path is changed. Availability of large amounts of parallel computing resources in the

form of "DOD Challenge Application Awards" has made it possible to carry out realistic simulations of ceramic materials processing and properties.

For the next generation of aerospace engines and high efficiency and environmentally clean turbines, it will be necessary to have materials that are mechanically stable at or above 1700°C. This is a very challenging problem. Synthesis, processing, and simulations will have to be carried out concurrently to accomplish the objective of making such materials.

In recent years, a great deal of progress has been made in the synthesis of ceramics that are much more ductile than conventional coarse-grained materials (Gleiter 1992; Siegel 1996). These so-called nanostructured materials are fabricated by *in situ* consolidation of nanometer size clusters. Despite a great deal of research, many perplexing questions concerning nanostructured ceramics remain unanswered. Experiments have yet to provide information regarding the morphology of pores or the structure and dynamics of atoms in nanostructured ceramics. As far as modeling is concerned, only a few atomistic simulations of nanostructured materials have been reported thus far. This is due to the fact that these simulations are highly computer-intensive: A realistic MD simulation of a nanostructured solid requires 10^5-10^6 time steps and ~ 10^6 atoms (each nanocluster itself consists of 10^3-10^4 atoms).

Interfaces

Roberto Car and his coworkers have performed many important materials simulations using the Car-Parrinello method (Car and Parrinello 1985). This interface simulation was carried out by Car (Princeton University, formerly Lausanne), A. Pasquarello (University of Geneva, Switzerland), and M.S. Hybertsen (Lucent Laboratories). This example reports a microscopic structural model of an interface between Si and its amorphous oxide generated by computer simulation (Pasquarello et al. 1998), as shown in Fig. 7.2. An atomic-scale characterization of the Si-SiO_2 interface is of great importance for silicon-based electronic devices, particularly in view of the requirement for increasingly thin (<50 Å) insulating oxide layers (Neaton et al. 2000). Unfortunately, in view of the disordered nature of the oxide, the available experimental information is limited to average properties such as the thickness and composition of the interfacial region. More detailed atomistic information can only come from theoretical modeling and simulation, which, in the present case, should accurately describe the chemical processes that cause structural rearrangement and ultimately give rise to the amorphous nature of the oxide.

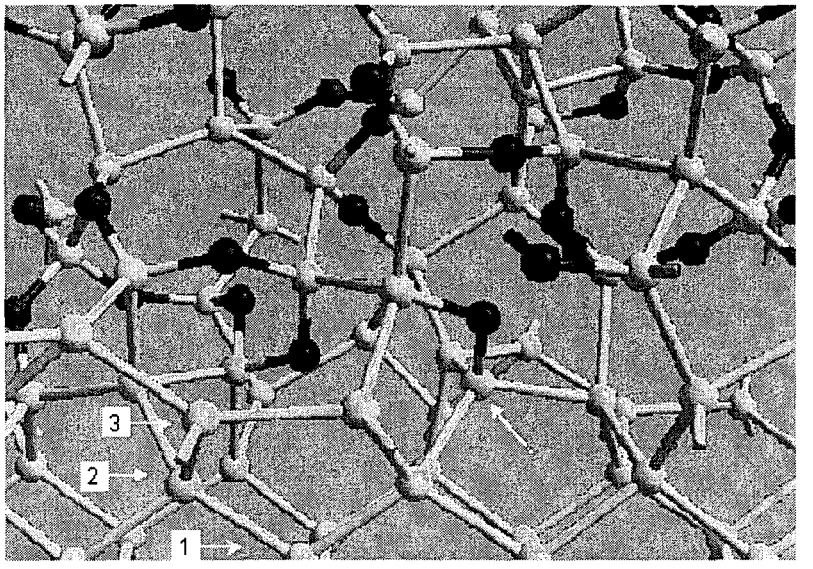


Fig. 7.2. Si(001)-SiO_2 interface structure from first-principles molecular dynamics. Oxygen atoms are dark gray (red in color versions of this report), silicon atoms are white (or yellow in color). Arrows 1, 2, and 3 indicate pure Si layers. Layers 1 and 2 are crystalline, layer 3 is disordered with Si dimers oriented along the interface. Excess Si atoms are present in this layer, consistent with ion scattering and x-ray reflectivity measurements (Pasquarello et al. 1998).

To realize this task the simulation described here was based on first-principles molecular dynamics (Car and Parrinello 1985). In this approach the electronic structure evolves self-consistently during the atomic motion and the atomic potential energy surface is calculated *on the fly* from the quantum ground state of the electrons within density functional theory. The results of the simulation, summarized in Fig. 7.2, were obtained with a code developed by R. Car and collaborators, using ultrasoft pseudopotentials and a plane-wave expansion of the electronic wavefunctions.

First-principles molecular dynamics methodologies have been applied successfully to a large variety of systems and materials (Parrinello 1997). At present the available computational resources limit these simulations to hundreds of atoms for time spans of a few tens of picoseconds. With the new computing power expected in the next decade, more realistic studies of oxidation and other important chemical processes will become possible, e.g., by embedding a quantum simulation of the reactive region into a classical simulation of the environment. The power of the method will be further enhanced by combining it with schemes that deal with rare events and activated processes, which have been limited so far mostly to classical molecular dynamics studies.

Carbon Nanostructures

Since their discovery in 1991, carbon nanotubes have been investigated vigorously for their exceptional mechanical and electronic properties. The nanotubes, sheets of carbon atoms rolled into cylinders a few billionth of a meter wide, are the stiffest and strongest nanoscale fibers known so far. NASA's interest in nanotubes derives from their possible use as reinforcing fibers in ultra lightweight multi-functional composites for aerospace structural applications (Srivastava et al. 2000).

Dr. Deepak Srivastava and his collaborators at NASA's Ames Research Center, California have simulated nanomechanics, plastic failure and functionalization chemistry of carbon nanotubes. The elastic response and plastic failure mechanisms of single-wall nanotubes were investigated with non-orthogonal tight-binding quantum molecular dynamics simulations and *ab initio* electronic structure methods. The elastic limit of thin carbon nanotubes, under axial compression, was found to be significantly lower than what was proposed earlier by classical molecular dynamics and continuum mechanics based descriptions (Srivastava et al. 1999). A novel mechanism of nanoplasticity of carbon nanotubes was observed in which bonding geometry collapses from a graphitic to a localized diamond-like reconstruction at the location of the collapse (Fig. 7.3). The computed critical stress (~153 GPa) and the locally collapsed structure were in close agreement with experimental observations of nanotubes in polymer composites. These results were the first to highlight the difference in the observed nanomechanics of nanotubes simulated within a quantum MD framework with the earlier classical atomistic descriptions.

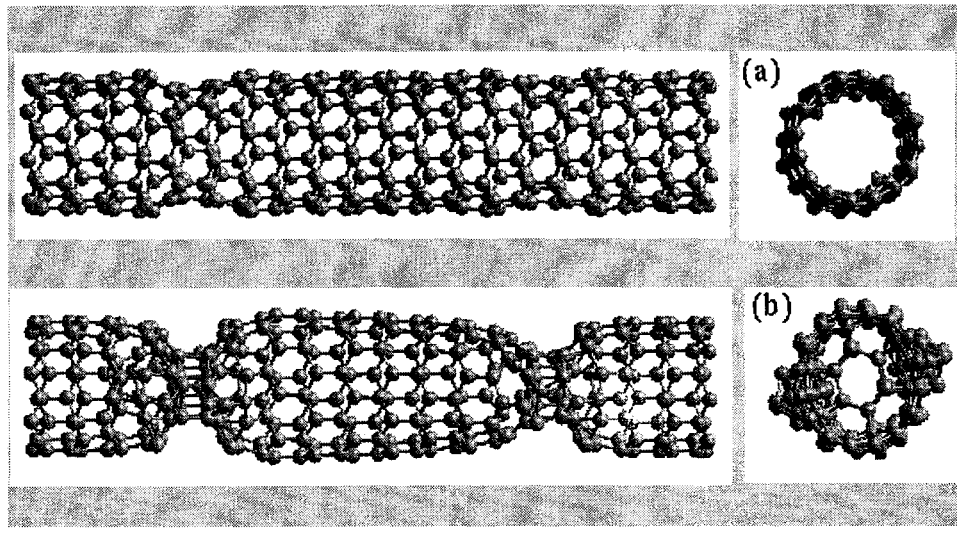


Fig. 7.3. An (8,0) C nanotube (a) before collapse, and (b) after collapse. The cross-sectional view shows formation of diamond like bonds at the location of the collapse (Srivastava et al. 2000).

FUTURE DIRECTIONS

Multiscale Systems

Processes such as fracture in real materials involve structures on many different length scales. They occur on a macroscopic scale but require atomic-level resolution in highly nonlinear regions. To study such multiscale materials processes, it is necessary to use a multiscale simulation approach that can describe physical and mechanical processes over several decades of length scales.

F.F. Abraham (IBM-San Jose), J.Q. Broughton (currently at J.P. Morgan, New York—at Naval Research Laboratory in Washington when the work was done), N. Bernstein (Naval Research Laboratory), and E. Kaxiras (Harvard University) have developed a hybrid simulation approach. It combines quantum mechanical (QM) calculations within the tight-binding approximation with large-scale molecular dynamics (MD) simulations embedded in a continuum, which is handled with a finite element (FE) approach based on linear elasticity (Abraham et al. 1998).

Such a multiscale FE/MD/QM simulation approach is illustrated in Fig. 7.4 for a material with a crack. The total system to be simulated is denoted as S_0. A subregion denoted as S_1 ($\subset S_0$) near the crack exhibited significant nonlinearity, and hence it was simulated atomistically, whereas the rest of the system, S_0 - S_1, was accurately described by the FE approach (Tadmor et al. 1996). In the region S_2 ($\subset S_1$) near the crack surfaces, bond breakage during fracture and chemical reactions due to environmental effects are important. To handle such chemical processes, QM calculations must be performed in S_2, while the subsystem S_1 - S_2 can be simulated with the classical MD method. Figure 7.4 also shows typical length scales covered by each of the FE, MD, and QM methods.

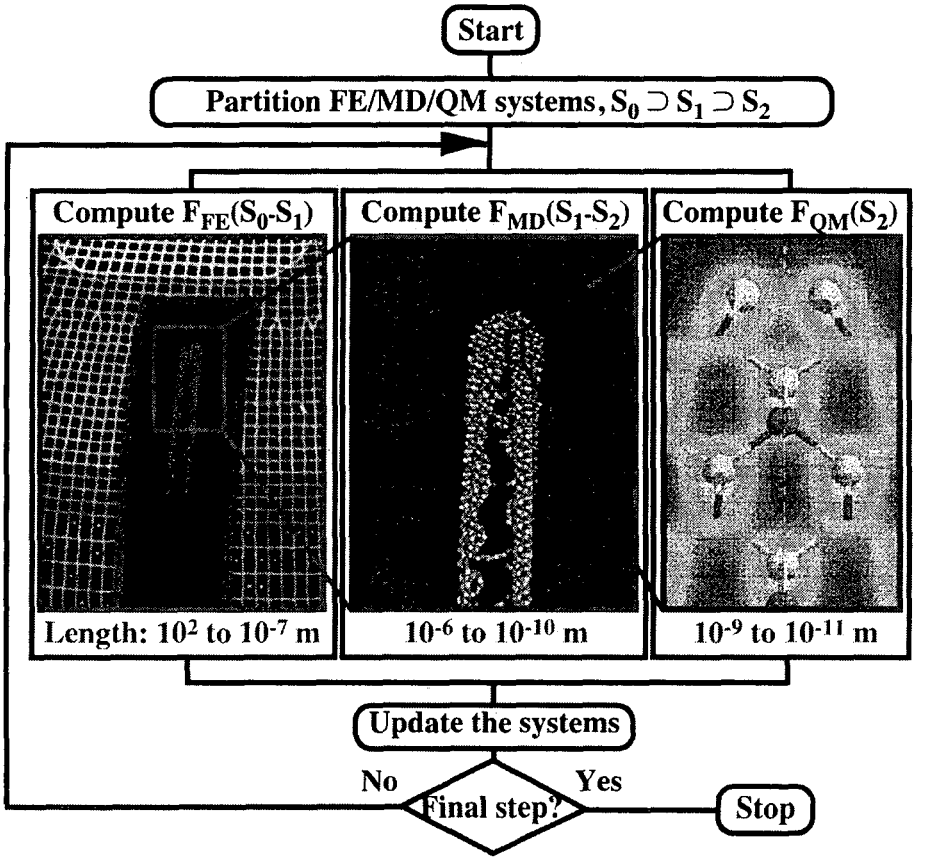


Fig. 7.4. Multiscale simulation approach methods (Abraham et al. 1998) combining the finite-element (left panel), molecular dynamics (middle panel), and *ab initio* electronic structure (right panel).

Moore's Law for Molecular Dynamics Simulations

Modern molecular dynamics simulations of materials started in 1964 when Aneesur Rahman simulated 864 argon atoms on a CDC 3600 computer in the Solid State Science Division of Argonne National Laboratory (Rahman 1964). The number of atoms that can be simulated in classical MD simulations doubled every 19 months to reach 8.1 billion atoms in 2000 (Fig. 7.5). Similarly, the number of atoms in DFT-MD simulations (started by Roberto Car and Michelle Parrinello in 1985 for 8 Si atoms) doubled every 13 months to 140,000 atoms in 2000.

Petaflop computers anticipated to be built in the next ten years should maintain the growth rates in these "MD Moore's Laws" (Shimojo et al. 2001), making it possible to perform 10^{12}-atom classical and 10^{7}-atom quantum MD simulations on such computers around year 2010. Multiresolution approaches used in these algorithms, combined with cache-conscious techniques, will be essential to achieve scalability on petaflop architectures. Ingenious use of multiscale FE/MD/QM simulations implemented with such scalable algorithms will play a significant role in materials research at the nanoscale.

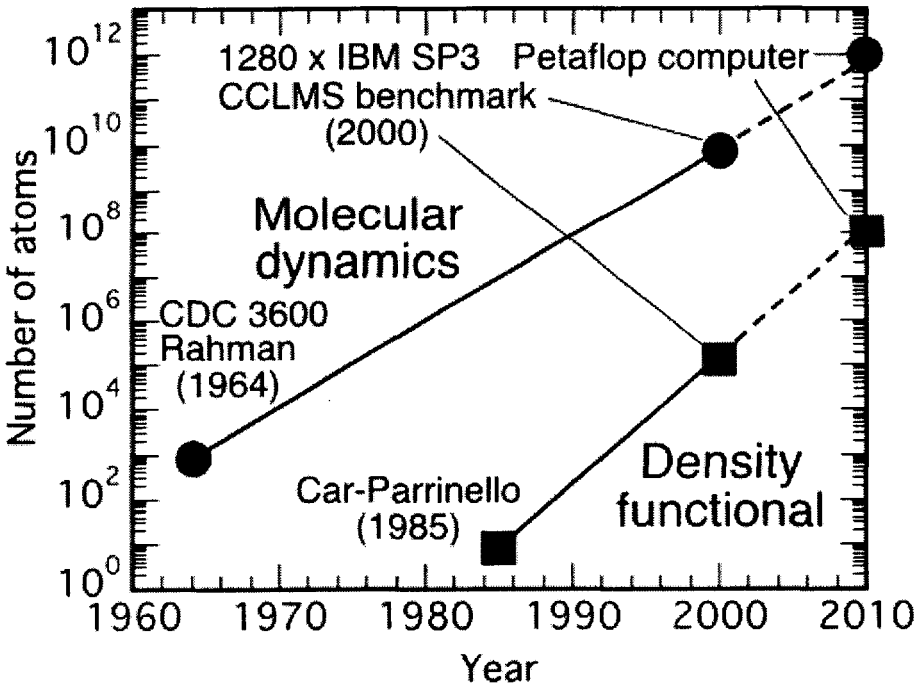


Fig. 7.5. "Moore's law" for molecular dynamics. The number of atoms in classical MD simulations (circles) has doubled every 18 months in the past 36 years, while that in quantum mechanical MD simulations (squares) has doubled every 13 months in the past 15 years. Petaflop computers anticipated to be built in the next ten years should maintain these rates.

However, there are still many problems to be overcome. For example, current linear-scaling DFT algorithms share serious problems, i.e., minimization is ill-conditioned (Goedecker 1999), therefore prohibiting their practical applications. This is due to the localization approximation that destroys the invariance of the energy under unitary transformations of the wave functions. Designing a scalable library of well-conditioned, fast FE/MD/QM algorithms that are applicable to a wide range of applications will be one of the most exciting challenges in computational science and engineering in the next ten years.

Grid Computing

"Grids" of geographically distributed supercomputing, mass-storage, and visualization resources (Fig. 7.6) connected via high-speed networks will revolutionize science and engineering by providing ubiquitous access to parallel computer simulation, data management, and visualization (Foster and Kesselman 1999). In order to use a "grid" as a single metacomputing platform for large-scale materials simulations, information

technology research is required, i.e., the development of adaptive management of heterogeneous grid resources; performance optimization on next-generation computer/network architectures; portable kernel libraries for adaptive, fast algorithms; and control theory for seamless embedding of hybrid software components.

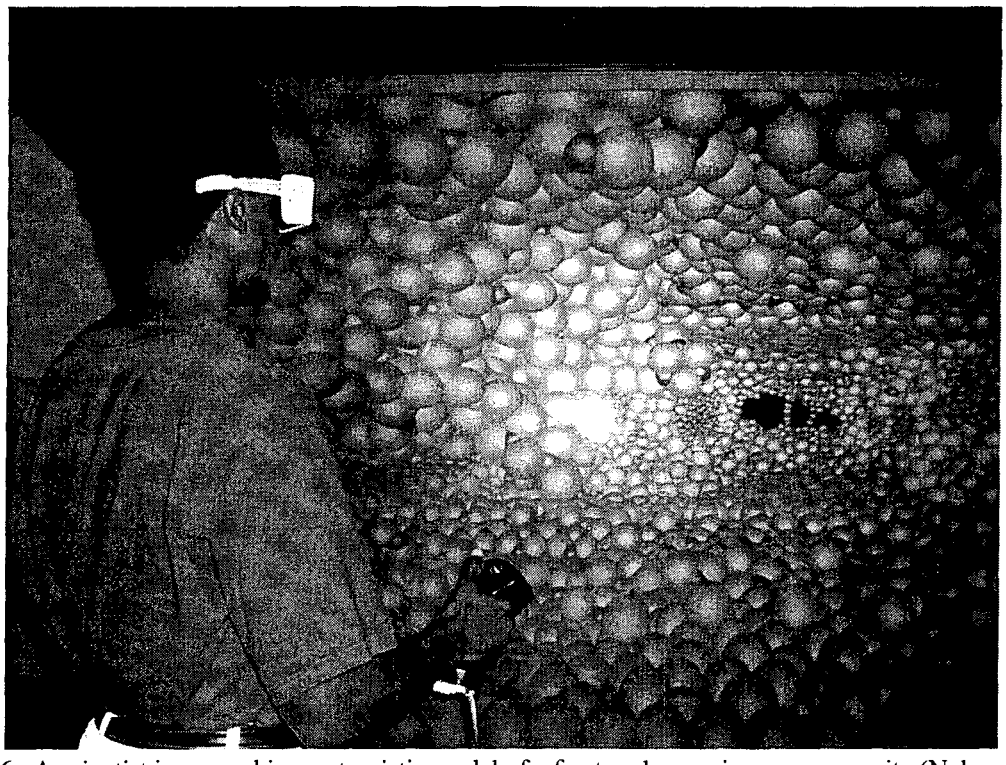


Fig. 7.6. A scientist immersed in an atomistic model of a fractured ceramic nanocomposite (Nakano et al. 1999). This system is a silicon nitride ceramic matrix reinforced with silica-coated silicon carbide fibers. Small spheres represent silicon atoms, and large spheres represent nitrogen (dark gray/green in color versions), carbon (darker gray/magenta), and oxygen (light gray/cyan) atoms (Nakano et al. 1999).

The large-scale multiscale simulations discussed above will generate enormous amounts of data. For example, a billion-atom MD simulation produces 100 gigabytes of data per frame including atomic species, positions, velocities, and stresses. Interactive exploration of such large datasets is important for identifying and tracking atomic features that are responsible for macroscopic phenomena. Immersive and interactive virtual environments are ideal platforms for such explorative visualization, as in Fig. 7.6. Scientists are designing algorithms to visualize large-scale atomistic simulations in three-dimensional immersive and interactive virtual environments. One approach employs adaptive and hierarchical data structures for efficient visibility culling and levels-of-detail control for fast rendering.

Bio-optoelectronics

Investigations are underway to explore whether it is possible to attach or embed semiconducting nanoclusters in biological systems such as DNA for biological computing applications. This kind of system has enormous potential applications in the field of electronics, biology, medicine, and biocomputing. In the near future, within 10 years or so, it may be possible to build bio-optoelectronic devices that may be embedded in microelectromechanical systems. This kind of bio-optoelectronic device will require advanced simulation methods, seamlessly incorporating classical MD and quantum calculations wrapped in continuum descriptions.

Expertise to carry out such complex simulations in the U.S. is available at universities (Harvard, North Carolina State, University of California at Berkeley, and Louisiana State), government laboratories (NASA-Ames and Sandia), and companies (Lucent, Motorola, and IBM). In Japan, these kinds of simulations will be

possible at the Institute of Solid State Physics and Tokyo Institute of Technology in Tokyo, and at the JRCAT and NEC research laboratories in Tsukuba. In Europe, Daresbury Laboratory and Cambridge University, UK, and Max Planck Institute in Stuttgart are the most likely institutions. High-end simulation research in UK industry appears to be weak or nonexistent. IBM-Zürich in Switzerland and Philips in the Netherlands are the promising companies in Europe for bio-optoelectronic research.

COMPARISON OF SCIENCE (MATERIALS SIMULATIONS OF METALS, SEMICONDUCTORS, CERAMICS, AND CARBON SYSTEMS) AND INDUSTRIAL APPLICATION OF SCIENCE IN THE UNITED STATES, EUROPE AND JAPAN

This section presents a comparative assessment of the current status and future trends of the science of materials modeling and its industrial applications. This is done for the United States, Europe and Japan. In addition, there is the issue of training a new generation of scientists and engineers in the *already-here-era* of information technology and nanoscience and technology. It is reasonable to ask how much of the information technology aspects of education should be included in the undergraduate and graduate curriculum at universities if academia is to produce successful simulation scientists and engineers for research and development in government laboratories and industry.

In addition to the material presented in this chapter, the reader should also refer to the site reports which resulted from the panel's visits to a large number of academic, government and industrial research organizations in the United States, Europe, and Japan. These reports provide rich source material on the current status of simulation research and its industrial applications, and a comparative assessment of the field as viewed from the vantage of the institutions visited.

Methods

Classical simulation methods and molecular dynamics. Modern molecular and materials simulation as we know it today began when Rahman wrote his seminal paper on MD simulation of liquid argon (Rahman 1964). Since then, many improvements have been made to the MD method. Some of the noteworthy developments include simpler and more efficient integration algorithms (Verlet, France); constraint MD algorithms (Berendsen, The Netherlands); constant-pressure MD (Anderson, U.S.); constant-temperature MD (Nose, and Nose and Klein, Canada); variable-shape MD box method (Parrinello and Rahman, U.S.); and symplectic multiple-time-step algorithms (Berne and collaborators, U.S.). To address a few quantum particles in a bath of classical particles, the Feynman path integral method has also been implemented using Monte Carlo (MC) and molecular dynamics (MD) methods.

Interatomic interaction potentials. One of the main limitations of classical simulation methods—MC and MD—is the unavailability of high-quality interatomic interaction potentials for a large variety of materials systems of interest. All early development and applications of MD and MC used Lennard-Jones two-body potential and, for simple ionic materials, simple two-body potentials were used which included long-range Coulomb interactions. Ewald summation method was used to deal with the long-range nature of Coulomb potential. The number of materials that could be simulated with these interaction potentials was quite limited.

Two developments in the area of interaction potentials stand out in late 1960s and early 1970s. These include interaction potentials developed by Stillinger and Rahman (U.S.) for simulating water and a broad effort to develop interaction potentials for biological systems by physical chemists. This was usually done in conjunction with constrained dynamics algorithms (Europe and U.S.). Interaction potentials for a wide variety of materials were also developed in UK at the Harwell Atomic Energy Establishment.

The most important development for dealing with covalent systems was made by Stillinger and Weber (U.S.) in 1985 when they proposed a simple three-body potential for silicon. This had a great impact on materials simulations because silicon is such an important material. Interaction potentials of similar form were also proposed for other covalent materials such as sulfur and selenium.

Charge-transfer interaction potentials have been developed by Vashishta and collaborators (U.S.) that include three-body bond bending and stretching to describe the covalent effects and long-range Coulomb interaction. These potentials take into account the charge transfer in materials such as SiO_2, $GeSe_2$, $SiSe_2$, Si_3N_4, SiC, Al_2O_3, GaAs, AlAs, and InAs. Bond-order potentials for a number of materials (Si, C, SiC) were proposed by Tersoff (U.S.). Most successful bond-order potentials for carbon-hydrogen system are due to Brenner (U.S.). Pettifor (UK) has also proposed a scheme for constructing bond-order potentials, but these potentials have not yet been tested on real materials. For metals, embedded-atom potentials proposed by Foiles, Baskes, and Daw (U.S.) have been quite successful.

Availability of high-quality interaction potentials remains one of the most serious needs for materials simulations. This is particularly true for industrial applications, where quick response is needed to "What if?" kinds of questions for materials properties under a variety of temperatures, pressures, and stress conditions. There seems to be no concerted program at universities or government laboratories to address the problem of developing interaction potentials for a wide variety of materials of interest. This is true in U.S., Europe, and Japan. This is an important problem and a serious limitation for classical MC or MD materials simulations—because on a per-atom basis, the Car-Parrinello method, which does not require interatomic potentials, is about 10 million times more computationally expensive.

Quantum mechanics. Lack of reliable interaction potentials for real materials for an arbitrary configuration of atoms has been the most serious limitation for classical MC and MD simulations. This problem was addressed through an innovative method proposed by Roberto Car and Michele Parrinello in 1985 (Europe). The method combined quantum mechanical description of electrons within the framework of density functional theory with classical dynamics of nuclei using molecular dynamics method. In the original version, simulated annealing method was proposed to solve for the electronic degrees of freedom. This method, while computationally very expensive, did not require interaction potentials between atoms/ions. Many important problems have been solved using the Car-Parrinello method. Also, a number of modifications and refinements to the method have been proposed in recent years.

A great deal of expertise with the Car-Parrinello method exists in Japan, especially at industrial research laboratories like NEC and Fujitsu. In Europe, Daresbury Laboratory and Cambridge University in UK have the broadest and most experienced groups in the Car-Parrinello method. In U.S., the expertise is located at several universities, at DOE laboratories like Livermore and Sandia, and at Motorola.

Scalable algorithms. For extending the scope of materials simulations to larger spatial and temporal scales, it is essential to develop efficient algorithms that utilize multi-resolutions in both space and time. Most of the developments in the area of scalable algorithms have taken place in the U.S. and, to a lesser extent, in Europe. Much less development has taken place in Japan.

General comparison. For methods development, it is fair to say that research and its applications are more advanced in the U.S. This is primarily because the molecular dynamics method (by Rahman) and density functional theory (by Kohn and collaborators) were discovered in 1964 in the United States. In addition, development in the 1980s and 1990s of parallel computing technology and *O(N)* algorithms, especially FMM for long range Coulomb forces, have kept the United States ahead of Europe and Japan. However, it should be emphasized that Europe and Japan are catching up very fast. In particular, note the site reports for Daresbury (UK), IBM-Zürich (Switzerland), TU-Denmark, ISSP (Japan), JRCAT-Tsukuba (Japan), NEC (Japan) and Fujitsu (Japan).

Applications

Metals. Dislocation dynamics represents a very important class of simulations. The main driving force for this effort in the United States is the ASCI program for nuclear weapons. At government laboratories such as Lawrence Livermore and Sandia National Laboratories, this kind of program is producing good results and a deeper understanding of experimental results. This kind of simulation will have wide ranging application in the design of electronic circuits where metal interconnects are used in micron and submicron dimensions, often at high temperatures and under stress. These kind of simulations will be carried out at companies like Intel, Motorola, IBM in the U.S.; NEC, Fujitsu, Toshiba in Japan; and Philips in Europe. A great deal of expertise for these kind of simulations already exists in Europe and Japan.

Semiconductors and semiconductor interfaces. This is and has been a very active research and development area in the United States and Japan. There is much less activity in Europe because of a general lack of high-end electronics and semiconductor industry.

In the research and development area leading to devices, the main companies active in the United States are Lucent Bell Laboratories, IBM (Yorktown and San Jose), Motorola, and Rockwell. During the last decade or so, Intel has focused its research mostly on the development of its main product line.

In Japan, semiconductor and semiconductor interfaces research, because it relates directly to the electronics and communications industries, is located at companies like NEC, Fujitsu, and Toshiba. A number of MITI laboratories in the Tsukuba area also have strong research programs. All universities have research programs in this area in departments of physics, applied physics and electrical and computer engineering. ISSP in Tokyo acts as a hub for a large part of the university-based research on these materials and phenomena.

Besides Philips in the Netherlands, there are no companies in Europe with the breadth and scope of research like NEC, Fujitsu, and Toshiba in Japan, and IBM, Lucent, and Motorola in the United States.

Optoelectronic devices, where GaAs plays an important role, are not yet major products, but they and their bio-optoelectronic forms will become more and more important.

Silicon carbide, on the other hand, is becoming increasingly more important because of its high-temperature electronic-device applications. Companies that have developed around Research Triangle Park and North Carolina State University are leading manufacturers of high purity SiC for devices. Japan also has strong companies in this area. No large international market for SiC has developed as yet. But with increasing need for electronic devices that will operate at high temperatures, SiC is becoming increasingly important.

Ceramic materials. This is an area where Japan has been the leader over many decades. Base materials for ceramic research are cheap; however, the research requires very painstaking, systematic and empirical studies. This is not revolutionary research but it requires enormous patience and persistence. Researchers in Japan excel in this kind of work because the attitudes it requires—patience, perseverance, and long-term commitment in spite of slow progress—are among the main virtues of Japanese culture. There are ceramic programs at universities, Tokyo Institute of Technology, government laboratories and companies. The most exciting areas of research are in nanophase ceramics and ceramic fiber composites. In Japan, many ceramics companies are located in the Nagoya-Osaka area.

There are only a few U.S. companies with strong ceramics research programs, e.g., Rockwell and Allied Signal (merged into Honeywell, which was purchased in 2000 by General Electric). Support for ceramic research is mainly provided by DOD agencies and DOE. Air Force, Navy and Army laboratories and DOE laboratories such as Argonne and Oak Ridge have strong ceramics programs.

Ceramics research is widespread in Europe. Germany has strong research groups at government laboratories.

Carbon systems. C_{60} and related structures were discovered in the United States at Rice University, and the carbon nanotube was discovered by Ijima at NEC in Japan. There is a broad-based research program in carbon systems in the United States at universities, government laboratories, and in industry. Because no revolutionary industrial applications of these systems have yet been found, the research is at a steady state. A number of companies and government laboratories are also interested in building tough and lightweight mechanical structures. In the United States, these include NASA and space-related industries. In Japan, there is also a broad research program at universities and substantial research programs at several companies. The main interest is focused toward applications of carbon nanotubes in nanoscale electronic devices. In the area of nanotubes, most of the materials simulation research in Japan and the U.S. is geared toward using suitably doped nanotubes as components in nanoscale electronic devices, field emitters for electron gun applications, and toward exploiting the unusually high mechanical strength of the nanotubes.

Much of this work in Japan goes on at the ISSP, among the best solid-state research institutions in Japan and in the world. The supercomputing facility at the institute provides parallel computing resources to all the university-based researchers in Japan. All computer codes used by researchers are their own or have been

developed with other university researchers. Users in this area of research are distributed at all major universities. The groups use all the methods of solid-state/condensed-matter physics. The first-principles calculation of material properties are carried out using a variety of electronic structure methods. These include: density functional theory, plane-wave basis sets, full-potential linearized augmented plane wave, linear muffin tin orbital, local spin density approximation, GW approximation, and new hybrid methods.

Future Directions

Multiscale simulations. There are processes in materials that involve structures at many different length scales. Fracture is an example of such a process. Even when one observes fracture in gigantic rocks and mountain formations, the small scale at which molecular bonds are broken is always there. However, a description of fracture in a material does not have to be carried out at atomistic level for all length scales. It is possible to design descriptions that are atomistic at smallest length scales, while at longer scales, a description of suitably lower resolution is chosen, the outermost description of the system being in terms of continuum elasticity. Like fracture, turbulence in fluids is another problem where a multiscale description is needed to describe the process.

During the last few years, efforts have been made to implement multiscale simulations for materials. It is believed that within the next 10 years, multiscale simulations will be possible in which quantum-mechanical description of bond formation and breaking (within density functional theory) will be embedded in a classical MD simulation, which will be wrapped in a continuum elasticity description using the finite element method. Because the quantum simulations are very time-consuming, even on the largest parallel computers available today, one can only carry out quantum simulations of only a few hundred atoms, surrounded by a few million classical atoms wrapped in a continuum elasticity model for only a few thousand steps. These methods are under development in several institutions. The main advantage of such methods will be that in the crucial region of the system where bonds are being broken and formed, there will be no need of a classical force field. These are the kind of simulations that will have the strongest impact on materials simulations.

Order N methods. Another possibility is to carry out the whole simulation by quantum methods. Because quantum simulations are very time-consuming (by a factor of 10^6–10^7 more compared to classical MD), it is unlikely that a very large fully quantum simulation will soon be possible on even the largest parallel computer. The most serious problem in this context is the scaling of system size; i.e., how does a quantum simulation scale with number of electrons? The scaling of Car-Parrinello methods is $O(N^3)$. In addition, the fast Fourier transform used in the method require global communications—not a very desirable thing if the problem is to be computed on a parallel machine with an increasingly larger number of processors. Robust, real space, *O(N)* algorithms for quantum simulations will have to be developed which are scalable and portable on a variety of parallel computers.

Grid computing approaches. Ultimately, large simulations will be implemented on a grid computer. It consists of geographically distributed supercomputing, mass-storage, and visualization resources connected via high-speed networks that will revolutionize science and engineering simulations. In order to use a grid-type computer as a single metacomputing platform for large-scale materials simulations, information technology research is required.

Other developments in computers. The cost of parallel computing is coming down drastically because of the availability of commodity hardware: PCs at around $1,000/box and various levels of interconnect hardware from $200/port to $2,000/port for gigabit or faster networks. Most of the software needed for such PC-based systems is available free of charge. Commercial simulation software is lacking at this time. It is hoped that commercial software will be available in versions that are truly scalable and portable from a single PC or workstation to a cluster for hundreds to a thousand.

In the United States, this transformation is already beginning at the universities and government labs and, to a lesser extent, in industry. Because the prices of commodity PC's and interconnect hardware are considerably higher in Europe and Japan, the transformation toward parallel computing has been slower. When the new generation of students who are considerably more comfortable with parallel computing join the industrial research groups, parallel materials simulations of real materials and processes will take a big step forward.

CONCLUSIONS

For the next few years, the United States will continue to lead in the field of materials simulations. This lead will diminish in time as the cost of computing goes down and as scalable and portable simulation software becomes readily available.

During our site visits, we noticed that Japanese computer companies like NEC and Fujitsu are seriously developing simulation software that will be sold separately or in conjunction with their hardware products. This could bring about a rapid advance in materials simulation efforts in Japan.

One of the major drawbacks of U.S. funding methods is that there is no concerted effort to develop interaction potentials for materials simulations. This is not considered an exciting area of research. However, without high-quality interaction potentials, it is not possible to carry out realistic materials simulations. Quantum simulation can bypass the need for interatomic potentials, but these simulations are 10^6 to 10^7 times more time consuming.

There is also no concerted effort for the development of simulation algorithms. The Information Technology Initiative may help alleviate this problem.

REFERENCES

Abraham, F.F. 1997. Portrait of a crack: rapid fracture mechanics using parallel molecular dynamics. *IEEE Comput. Sci. Eng.* 4 (2), 66.

Abraham, F.F., J.Q. Broughton, N. Bernstein, and E. Kaxiras. 1998. Spanning the length scales in dynamic simulation. *Comput. Phys.* 12, 538.

Allen, M.P., and D.J. Tildesley. 1987. *Computer Simulation of Liquids.* Oxford Univ. Press, Oxford.

Bazant, M.Z., E. Kaxiras, and J.F. Justo. 1997. Environment-dependent interatomic potential for bulk silicon. *Phys. Rev. B* 56, 8542.

Brenner, D.W. 2000. The art and science of an analytic potential. *Phys. Stat. Sol. (b)* 217, 23.

Campbell, T.J., R.K. Kalia, A. Nakano, P. Vashishta, S. Ogata, and S. Rodgers. 1999. Dynamics of oxidation of aluminum nanoclusters using variable charge molecular-dynamics simulations on parallel computers. *Phys. Rev. Lett.* 82, 4866.

Car, R., and M. Parrinello. 1985. Unified approach for molecular dynamics and density functional theory. *Phys. Rev. Lett.* 55, 2471.

Dapprich, S., I. Komáromi, K.S. Byun, K. Morokuma, and M.J. Frisch. 1999. A new ONIOM implementation in Gaussian 98. I. The calculation of energies, gradients, vibrational frequencies, and electric field derivatives. *J. Mol. Struct. (Theochem)* 461-462, 1.

de la Rubia, T.D., H.M. Zhih, and M.J. Caturia. 2000. Multiscale modelling of plastic flow localization in irradiated materials. *Nature* 406, 871.

Deuflhard, P., J. Hermans, B. Leimkuhler, A.E. Mark, S. Reich, and R.D. Skeel. 1999. *Computational Molecular Dynamics: Challenges, Methods, Ideas.* Springer, Berlin.

Eichinger, M., P. Tavan, J. Hutter, and M. Parinello. 1999. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.* 110, 10452.

Field, M.J., P.A. Bash, and M. Karplus. 1990. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* 11, 700.

Foiles, S.M., M.I. Baskes, and M.S. Daw. 1986. Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. *Phys. Rev. B* 33, 7983.

Foster, I., and C. Kesselman. 1999. *The Grid: Blueprint for a New Computing Infrastructure.* Morgan Kaufmann, San Francisco.

Fox, G.C., R.D. Williams, and P.C. Messina. 1994. *Parallel Computing Works.* Morgan Kaufmann, San Francisco, CA,.

Germann, T.C., and P.S. Lomdahl. 1999. Recent advances in large-scale atomistic materials simulations. *Comput. Sci. Eng.* 1 (2), 10.

Gleiter, H. 1992. Materials with ultrafine microstructures: retrospectives and perspectives. *Nanostructured Mater.* 1, 1.

Goedecker, S. 1999. Linear scaling electronic structure methods. *Rev. Mod. Phys.* 71, 1085 (1999).

Greengard, L., and V. Rokhlin. 1987. A fast algorithm for particle simulations. *J. Comput. Phys.* 73, 325.

Hohenberg, P., and W. Kohn. 1964. Inhomogeneous electron gas. *Phys. Rev.* 136, B864.

Kalia, R.K., T.J. Campbell, A. Chatterjee, A. Nakano, P. Vashishta, and S. Ogata. 2000. Multiresolution algorithms for massively parallel molecular dynamics simulations of nanostructured materials. *Comput. Phys. Commun.* 128, 245.

Kohn, W., and P. Vashishta. 1983. General density functional theory. In *Inhomogeneous Electron Gas*, eds. N.H. March and S. Lundqvist. Plenum, New York, p. 79.

Martyna, G.J., M.E. Tuckerman, D.J. Tobias, and M.L. Klein. 1994. Explicit reversible integrators for extended systems dynamics. *J. Chem. Phys.* 101, 4177.

Nakano, A. 1997. Fuzzy clustering approach to hierarchical molecular dynamics simulation of multiscale materials phenomena. *Comput. Phys. Commun.* 105, 139.

Nakano, A. 1999. Multiresolution load balancing in curved space: the wavelet representation. *Concurrency: Practice and Experience* 11, 343.

Nakano, A., R.K. Kalia, and P. Vashishta. 1994. Multiresolution molecular dynamics algorithm for realistic materials modeling on parallel computers. *Comput. Phys. Commun.* 83, 197.

Nakano, A., R.K. Kalia, and P. Vashishta. 1999. Scalable molecular-dynamics, visualization, and data-management algorithms for materials simulations. *Comput. Sci. Eng.* 1 (5), 39.

Ogata, S., H. Iyetomi, K. Tsuruta, F. Shimojo, R.K. Kalia, A. Nakano, and P. Vashishta. 1999. Variable-charge interatomic potentials for molecular-dynamics simulations of TiO_2. *J. Appl. Phys.* 86, 3036.

Ogata, S., F. Shimojo, A. Nakano, P. Vashishta, and R.K. Kalia. 2001. Hybrid quantum mechanical/molecular dynamics simulation on parallel computers: density functional theory on real-space multigrids. *Comput. Phys. Commun.*, in press.

Omeltchenko, A., T.J. Campbell, R.K. Kalia, X. Liu, A. Nakano, and P. Vashishta. 2000. Scalable I/O of large-scale molecular-dynamics simulations: a data-compression algorithm. *Comput. Phys. Commun.* 131, 78.

Parrinello, M. 1997. From silicon to RNA: the coming of age of ab initio molecular dynamics. *Solid State Commun.* 102, 107.

Pasquarello, A., M.S. Hybertsen, and R. Car. 1998. Interface structure between silicon and its oxide by first-principles molecular dynamics. *Nature* 396, 58.

Payne, M.C., M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos. 1992. Iterative minimization techniques for *ab initio* total energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.* 64, 1045.

Pechenik, A., R.K. Kalia, and P. Vashishta. 1999. *Computer-Aided Design of High-Temperature Materials.* Oxford Univ. Press, Oxford.

Pettifor, D.G., and I.I. Oleinik. 2000. Bounded analytic bond-order potentials for sigma and pi bonds. *Phys. Rev. Lett.* 84, 4124.

Rahman, A. 1964. Correlations in the motion of atoms in liquid argon. *Phys. Rev.* 136, A405.

Rapaport, D.C. 1995. *The Art of Molecular Dynamics Simulation.* Cambridge Univ. Press, Cambridge.

Rappé, A.K., and W.A. Goddard. 1991. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* 95, 3355.

Rick, S.W., S.J. Stuart, and B.J. Berne. 1994. Dynamical fluctuating charge force field: application to liquid water. *J. Chem. Phys.* 101, 6141.

Schlick, T., E. Barth, and M. Mandziuk. 1997. Biomolecular dynamics at long timesteps: bridging the timescale gap between simulation and experiment. *Annual Rev. Biomol. Struct.* 26, 181.

Shimojo, F., T.J. Campbell, R.K. Kalia, A. Nakano, P. Vashishta, S. Ogata, and K. Tsuruta. 2001. A scalable molecular-dynamics-algorithm suite for materials simulations: design-space diagram on 1,024 Cray T3E processors. *Future Generation Comput. Sys.*, in press.

Siegel, R.W. 1996. Creating nanophase materials. *Scientific American*, Dec., p. 74.

Singh U.C., and P.A. Kollman. 1986. A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems. *J. Comput. Chem.* 7, 718.

Skeel, R.D., G. Zhang, and T. Schlick. 1997. A family of symplectic integrators. *SIAM J. Sci. Comput.* 18, 203.

Smit, B., and D. Frenkel. 1996. *Understanding Molecular Simulation: From Algorithms to Applications.* Academic Press.

Srivastava, D., et al. 2000. Carbon-nanotube-based nanotechnology in an integrated modeling and simulation environment. Chapter 14, *Handbook of Nanostructured Materials and Nanotechnology,* ed. H.S. Nalwa, Vol. 2, 665. Academic Press.

Srivastava, D., M. Menon and K. Cho. 1999. Nanoplasticity of single-wall carbon nanotubes under uniaxial compression. *Phys. Rev. Lett.* 83, 2973.

Stillinger, F., and T.A. Weber. 1985. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* 31, 5262.

Strauch, D., and B. Dorner. 1990. *J. Phys.: Condens. Matter* 2, 1457.

Streitz, F.H., and J.W. Mintmire. 1994. Electrostatic potentials for metal-oxide surfaces and interfaces. *Phys. Rev. B* 50, 11996.

Tadmor, E.B., M. Ortiz and R. Phillips. 1996. Quasicontinuum analysis of defects in solids. *Philos. Mag. A* 73, 1529.

Tersoff, J. 1986. New empirical model for the structural properties of silicon. *Phys. Rev. Lett.* 56, 632.

Toukmaji, A.Y., and J.A. Board. 1996. Ewald summation techniques in perspective: a survey. *Comput. Phys. Commun.* 95, 73.

Troullier, N., and J.L. Martins. 1991. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* 43, 1993.

Udron, D., et al. 1991. *J. Non-Cryst. Solids* 137&138, 131.

Vashishta, P., R.K. Kalia, and A. Nakano. 1999. Large-scale atomistic simulation of dynamic fracture. *Comput. Sci. Eng.* 1 (5), 56.

Vashishta, P., R.K. Kalia, A. Nakano, W. Li, and I. Ebbsjö. 1996. Molecular dynamics methods and large-scale simulations of amorphous materials. In *Amorphous Insulators and Semiconductors*, eds. M.F. Thorpe and M.I. Mitkova. NATO ASI.

Wang, C.Z., and K.M. Ho. 1996. Tight-binding molecular dynamics study of covalent systems. In *Advances in Chemical Physics Vol. XCIII*, eds. I. Prigogine and S.A. Rive. Wiley, New York.

CHAPTER 8

APPLICATIONS: CATALYSIS

Matthew Neurock

INTRODUCTION

Catalysis is at the heart of nearly all chemical manufacturing processes and is responsible for over $400 billion in annual revenue in petroleum refining and chemical manufacturing (Bell et al. 1992; Hagen 1999). Over 90% of all chemical processes involve some form of catalytic transformation step. Nearly 60% of all of the chemical intermediates manufactured are made via catalytic routes. This includes 40 of the 50 largest commodity chemicals (Bell et al. 1992; Farrauto and Bartholomew 1997).

In addition to its tremendous economic impact, catalysis also plays a critical role in the design of environmentally green chemical processes. The design of more active and selective catalytic materials could significantly reduce (and in some cases even eliminate) the following:

- Waste by-product production
- Energy consumption
- Environmentally unfriendly solvents

It has been estimated that over 1.5 billion metric tons of carbon were released to the atmosphere in 1995 (Bell et al. 1992). At least 10% of this was due to automotive exhaust, stationary power sources and chemical manufacturing. The design of a more selective catalyst for any one of these processes could dramatically reduce these emissions.

The second leading environmental concern involves the escalating energy demands by most chemical processes. The design of more active catalytic materials that operate at much lower temperatures could dramatically cut back on energy consumption. Ammonia synthesis is run at temperatures on the order of 400°C in order to activate the N_2 bond (which is quite stable) and remove the NH_3 products from the surface. There is a great deal of research activity that is being carried out around the world aimed at designing an active catalyst that will lower the operating temperature for this process.

The final environmental issue involves eliminating the need for hazardous chemical solvents that ultimately become waste. Many of the acid-catalyzed chemistries are currently carried out in highly corrosive and environmentally unfriendly acids such as sulfuric acid and hydrogen fluoride. Novel solid acids, with acidities that approach or exceed these corrosive liquid-phase acids, are being synthesized and examined as suitable replacements for the liquid-phase acids.

Catalysis is quite pervasive and cuts across many different manufacturing industries, including sectors from petroleum, chemical, pharmaceuticals, energy and automotive. It is quite clear that catalysis will continue to impact our economy and take on a pivotal role in the design of green chemical processes. Most of the research that is currently carried out in industry, government and academic laboratories is focused on the

design of more active and highly selective materials. Material stability and lifetime are also considerable issues, especially for industrial laboratories. Catalytic kinetics becomes quite relevant as industries move toward scale-up. The primary goal of most of the industrial laboratories is therefore to accelerate catalyst development and scale-up.

There have been a number of outstanding advances in catalyst characterization, catalyst synthesis, surface science, transient kinetic studies and reaction engineering that have significantly improved our understanding of fundamental factors that control catalytic performance and have aided the design of new materials. Molecular modeling is rapidly developing as an excellent complement to many of these efforts. It is currently used in a variety of different ways to enhance both catalyst design and reactor scale-up activities. A recent NSF/DOE report, *Future Directions for Theoretical Catalysis, Homogeneous, Heterogeneous, and Surface Reactivity* (Neurock and Bell 1997), assesses the state of the art in molecular modeling applications to catalysis as of 1997. Much of the work prior to 1995 was directed towards understanding the chemistry, or what can be called "analysis." Many of the current studies, however, are beginning to expand these efforts to aid experimental research in identifying new materials, i.e., "design."

Nearly all of the chemical companies that were visited in this study had an active effort in molecular modeling of catalysis. As discussed below, most of this energy was concentrated on homogeneous catalysis, whereby the atomic structure of the catalyst is well defined. While there was clearly a very strong interest in future capabilities for modeling heterogeneous systems, most of the companies we spoke with had not ventured too far into modeling heterogeneous catalytic systems. A few of them, however, have actually begun some work in the area of heterogeneous catalysis.

In terms of applications, a number of important areas were identified as critical. Most of these areas were consistent with those identified in the U.S. Department of Energy's Vision 2020 roadmap for the future of catalysis (http://www.oit.gov/chemicals/page9a.shtml and Haynes et al. 2000) as listed in Table 8.1. Additional areas that were not specifically identified in the catalysis workshop report were identified in our current study. In addition to outlining the critical areas of chemistry and technology, the Vision 2020 Catalysis roadmap also identified a set of tools that will be necessary in the search for new materials. The combination of *in situ* experimental tools along with molecular modeling was identified at the top of the list (Haynes et al. 2000).

Table 8.1
Critical Areas for Molecularly Based Modeling of Catalysis

Areas cited in Vision 2020 report (Haynes et al. 2000)	Additional areas identified in this study
Selective oxidation	Fuel cell technology
Alkane activation	Enzyme catalysis
Olefin polymerization	Biomimetic materials
Selective stereo- and regio-synthesis	High-throughput screening
By-product and waste minimization	NOx reduction technology
Renewable feedstocks	

Catalysis can be broadly classified as either *homogeneous* or *heterogeneous*. A homogeneous catalyst is one that is soluble in the reaction medium. The chemistry is carried out in a single solution phase. Heterogeneous catalysts, on the other hand, are solid-state materials whereby the reactivity occurs at the liquid/solid or vapor/solid interface (or interphase). The discovery of single-site metallocene catalysts in the early 1990s brought about a resurgence in the research efforts for homogeneous catalysts. Metallocenes offer the hope of significant improvements over traditional Ziegler-Natta catalysts along with the ability to begin to tune the resulting polymer properties. As will be discussed later, much of the current efforts and successes of molecular modeling in catalysis have been on homogeneous metallocene catalysts.

Metallocenes are single-site catalysts that are made up of a metal center along with a sequence of ligands. While Zr, Ti, and Hf are most readily identified, a number of other metals have also proven to be active for specific steps. They were first identified as a more general class of organometallic complexes represented by

ferrocene, which contains a metal Fe atom center sandwiched between two cyclopentadienyl ligands. Metallocenes can contain various different types of ligands but typically have either one or two cyclopentadienyl type ligands. The metal atom center along with its ligands can be tuned in order to create unique electronic or steric features at the active center. These features ultimately control the stereochemistry at the active site, the monomer produced, the polymer product distribution, and the polymer microstructure (Fig. 8.1). This can enable the explicit design of physical and mechanical properties for the final polymer. Mechanistically, this mimics the "lock and key" chemistry of enzymes.

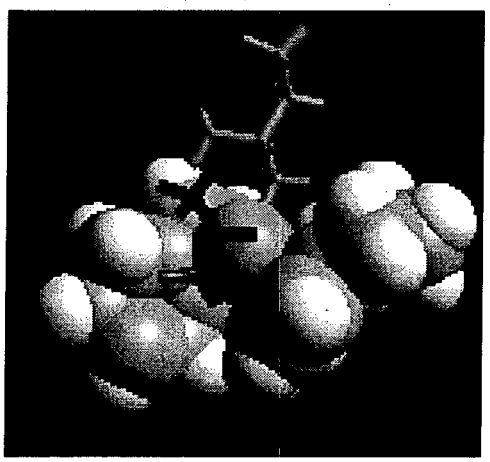


Fig. 8.1. Illustration of single-site metallocene complex (MSI, Inc.).

Metallocenes are currently used to produce polyolefin homopolymers and copolymers as well as a series of enantioselective reactions. They have had a tremendous impact in the manufacturing of polyolefins as well as other stereoselective polymers. It has been estimated that chemical manufacturers have spent well over one billion dollars in research toward understanding and developing metallocene catalysts. There is speculation that these efforts will increase significantly over the next decade as the production of thermal plastics and elastomers increases. In addition, metallocenes will also likely be used in the future for the synthesis of polymers with various polar functional groups.

Heterogeneous catalysts are made up of a variety of different materials including supported metals, metal oxides, zeolites, metal sulfides, metal carbides, metal nitrides and a number of other novel systems. This spans a very diverse range of solid-state materials that can all have very different properties and control catalytic chemistry by different operative mechanisms. Representative crystalline structures for several of these classes of materials are presented in Fig. 8.2. Zeolites are quite unique in that they typically have a well-defined atomic structure that can lead to changes in both chemical as well as physical processes. Their well-defined crystalline structures make them ideal for computational studies. Most of the other materials typically have very complex surface structures that require help not only from theory but also from analytical characterization in order to elucidate the possible surface structures and reactivities. For example, the chemistry on supported metals can be markedly affected by the metal particle size as well as metal dispersion. What controls this behavior is not necessarily the particle size but the number and nature of the edge defect sites and the surface facets that are exposed. *In-situ*, and in many cases even the *ex-situ*, resolution of these features is currently not possible from most characterization techniques.

METHODS

Molecular modeling of catalysis can roughly be divided into quantum chemical methods and atomic or molecular level simulations. Different techniques are used based upon the questions one wishes to answer. Any study aimed at extracting data on catalytic properties such as reactivity requires quantum chemical methods. There is, however, a large class of problems that deals with intra- and intermolecular transport phenomena. We use this distinction between quantum chemical methods and molecular simulation to categorize some of the methods that are currently employed to examine catalytic systems.

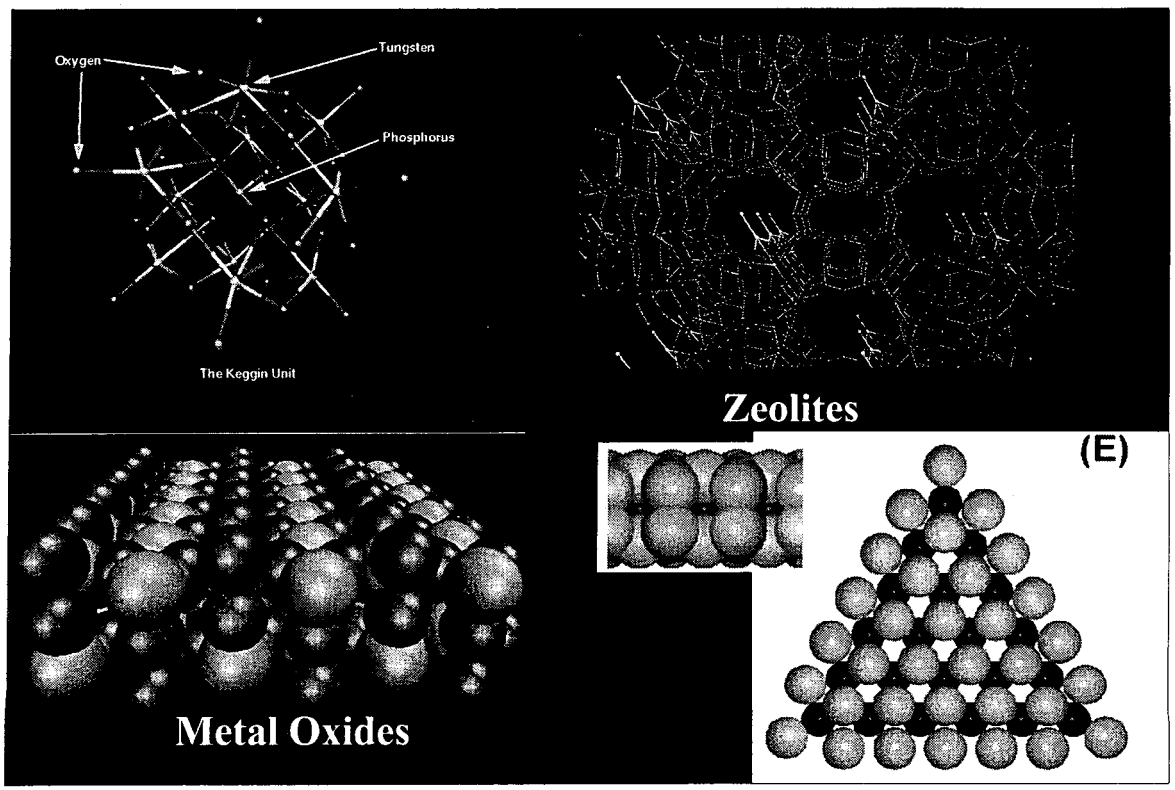


Fig. 8.2. Structure of different forms of crystalline catalytic materials: (upper left) the primary keggin unit of the heteropolyacids; (upper right) mordenite, a zeolite; (lower left) metal oxides: RuO_2 anatase; and (lower right) MoS_2 metal sulfides.

Quantum Chemical Methods

The application of quantum chemical methods to modeling heterogeneous catalysis can be categorized by the QM methods used as well as the structural models employed to describe the active surface structure. A full range of approaches exists for the solution of Schrödinger's wave equation. These include approximate methods such as Extended Hückel theory, semi-empirical methods, and *ab initio* methods. (These methods were discussed in detail in Chapter 2.) In addition there are a number of excellent reviews of these approaches and their application to surface bonding (van Santen and Neurock 1995; van Santen and Neurock 1998; Whitten and Yang 1996; Pacchioni 1995; Hammer and Norskov 2000; Hammer and Norskov 1997; Stampfl et al. 2001; Neurock 1997). Herein we discuss only the application of first-principles-based methods to catalysis. These methods can roughly be divided into molecular orbital and density functional theoretical (DFT) approaches. Both methods attempt to solve the Schrödinger equation.

Many of the earliest calculations were performed using Hartree-Fock theory. Hartree-Fock (HF) theory, however, is inadequate for predicting the properties of transition metal systems. More reliable treatments of electron-correlation along with multi-reference basis functions are essential in order to derive any type of quantitative effort. Most of these approaches, however, are not amenable to the large numbers of atoms necessary to account for active sites. Both configuration interaction (CI) and MP2 calculations have been used to examine model clusters with accuracies that are on the order of 5-10 kcal/mol for the calculated energies. MP2 scales as N^5 where N is the number of basis functions or, roughly speaking, the number of electrons. MP2 calculations can become prohibitive for treating larger metal clusters or surfaces. The use of effective core potentials to describe the core electrons has helped to considerably reduce the cost of these calculations whereby one can examine on the order of 10-20 metal atoms.

Density functional theory (DFT), as discussed in Chapter 2, is characteristically different in that it describes the energy of the ground state of an N-electron system as a direct functional of the density. Density functional methods scale at the level of N^3. This is significantly less than the N^5 dependence for MP2 and CI

methods. DFT has therefore been the clear choice by most of the industrial modelers who are after accurate and expedient methods. DFT predicted adsorption energies for various different adsorbates on different metal surfaces are compared against known experimental results in Fig. 8.3(a). The error bars are approximately 5-7 kcal/mol. Structural predictions are usually within 0.05 Å whereas spectroscopic properties are usually within about 5% of the experimental measures, as shown in Fig. 8.3(b).

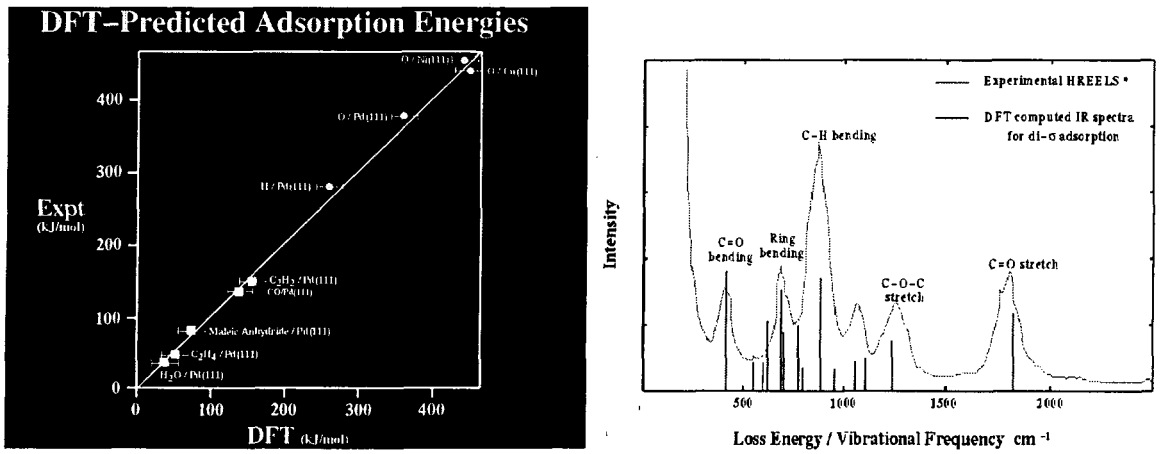


Fig. 8.3. Comparison of calculated and experimental (a, left) chemisorption energies and (b, right) vibrational frequencies for surface adsorbates (Venkataraman, Neurock, and Coulston 1999).

Although DFT has proven to be applicable across a broad range of chemistry, there are still areas where DFT has its difficulties. The prediction of weak van der Waals interactions and the properties of excited states are known issues for DFT. The latest DFT methods, which attempt to solve the time-dependent problem, may increase our abilities to predict these systems. Even with the latest advances there are situations where DFT will lead to errors outside the accuracies cited above.

Catalyst Models

The success in modeling catalytic systems not only depends upon the accuracy of the methods employed but also on the reality of the model system that is studied. There are three different techniques that are currently used to model the structure at the active site. These are known as the cluster (van Santen and Neurock 1995; van Santen and Neurock 1998; van Santen and Kramer 1995; van Santen 1997), embedded-cluster (Whitten and Yang 1996; Sauer 1989; Sauer 1994; Sauer et al. 1994; Sierka and Sauer 2000), and periodic methods (Hammer and Norskov 2000; Payne et al. 1992), respectively. Each of these methods has its own set of advantages and disadvantages. Characteristic models for each of these systems are presented in Figure 8.4.

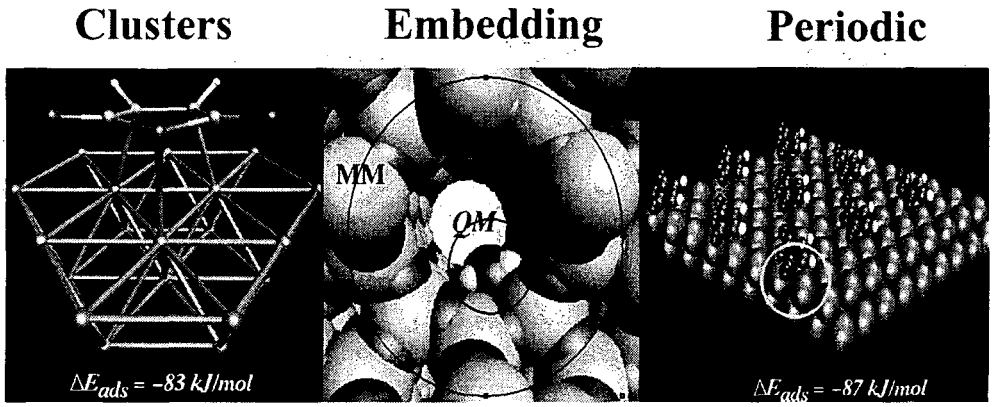


Fig. 8.4. Three approaches and examples for modeling chemisorption and reactivity on surfaces: cluster approach, maleic anhydride on Pd_{19} (left); embedding scheme—ammonia adsorption in a zeolite cage (center); periodic slab model: maleic anhydride adsorption on Pd(111) (right).

Cluster Approach

Cluster methods use a discrete number of atoms to model the active site region. The basic premise behind the approach is that chemisorption and reactivity are local phenomena, primarily affected by the nearby surface structure. Rigorous quantum mechanical calculations are therefore used to elucidate the extent of orbital overlap, Pauli repulsion, and electron correlation of the adsorbate/surface ensemble in order to predict adsorption geometry, adsorption energies and surface reactivity. The clusters can adopt different sizes and configurations depending upon the specific system being examined and the accuracy of the desired results.

An important advantage of the cluster approach is that the active site is explicitly described by the interactions between the local molecular orbitals of the adsorbate and the surface. Chemical intuition as regarding orbital interactions and Pauli repulsion is therefore direct. This enables one to easily probe bonding and reactivity and to formulate structure-property relationships. Well-established quantum mechanics and DFT algorithms can be used here in a routine manner. This allows one to take advantage of current computational algorithms that employ sophisticated and robust geometry optimization schemes and transition-state search methods.

A major limitation of the cluster approach is its incomplete representation of the electronic system provided by its small size and the discrete nature of the cluster employed (van Santen and Neurock 1995; Whitten and Yang 1996). The electronic structure of a catalytic system in the cluster approach is modeled by molecular orbitals, separated by discrete energy levels rather than by continuous valence and conduction band structures that describe the solid state. This leads to size-dependent phenomena that also exist in experiments on small metal clusters. Despite this approximation, the cluster approach has been used successfully to model a number of different catalytic systems. By carefully optimizing the adsorbate-cluster interactions, the lowest energy structures and electronic states can be calculated and used to predict reliable energetic data. A point worth noting is that many industrial catalytic systems involve small metal particles that are highly dispersed on a support, so they may behave more like a cluster rather than the bulk metal. Thus, the cluster approach ought to provide a reasonable starting point for these systems.

The cluster approach has also been used in the literature to model zeolites (van Santen and Kramer 1995; Sauer 1994; Sauer et al. 1994), metal oxides (Hermann, Witko and Michalak 1999; Witko, Herman and Tokarz 1999), and metal sulfides (Neurock and van Santen 1994). It is typically easier to model the electronic structures in these systems because of the discrete gap between their conduction and valence bands. Moderate-sized clusters usually provide reasonable representations for these systems. However, the two major challenges that arise involve cluster termination and long-range electrostatic interactions. Cluster termination has been examined in great detail for zeolitic systems. A widely accepted approach involves termination with hydroxyl groups, whereby the hydrogens are fixed at lattice positions that anchor the cluster to the actual zeolite (van Santen and Kramer 1995; Sauer 1994; Sauer et al. 1994; Nicholas 1999; Gonzales, Bell and Chakraborty 1997; Frash and van Santen 1999). The long-range Madelung interactions can be accounted for by introducing electric field effects into the Hamiltonian (Teunissen et al. 1994).

Embedded-cluster Approach

The embedded-cluster model is simply an extension of the cluster approach that attempts to treat the problems associated with the abrupt cluster termination. In the embedded-cluster approach, a rigorous QM method is used to model the local region about the active site. This primary cluster is then embedded in a much larger system in order to simulate the external electronic environment. The outer model employs a much simpler quantum mechanical treatment or an empirical force field to simulate the external environment. This minimizes cluster-size artifacts.

The difficulties with this approach involve accurately matching the electronic structure at the interface between the inner-cluster and the external model. The external model, which is derived from a much simpler method, can also be called into question. The approach, however, has been successfully used to model metals, zeolites, and metal oxides. This can require defining an interfacial region whereby the inner core model and the outer core are connected.

The newest methods treat the core region with DFT and the outer shell with molecular mechanics and can be carried out on systems that contain thousands to tens of thousands of atoms. They have been quite successful in modeling drugs, proteins and other biologically important systems.

Whitten (Whitten and Yang 1996) and Carter (Govind, Wang and Carter 1999) have used *ab initio* core regions in order to provide accurate description of the local bond breaking chemistry and simpler QM calculations to describe the exterior. Sauer (Sierka and Sauer 2000) has developed a general code that will use standard *ab initio* methods to treat the active region of a zeolite along with shell models to describe the oxide environment. Morokuma has developed a general approach, termed the ONIOM method (Froese and Morokuma 1999; Vreven and Morokuma 2000), which is currently offered as a standard tool in Gaussian 98 (Gaussian Inc.). It enables users to choose their favorite methods for the core and exterior regions.

Periodic Quantum-mechanical Calculations

The extended electronic structure and external field effects for a well-defined surfaces (metal, oxide, sulfide, other) can be modeled by using periodic infinite-slab calculations (Payne et al. 1992). In the periodic approach, a unit cell is defined to represent the system. This unit cell is then subsequently repeated in one, two, or three directions, thus providing the electronic structure for linear, slab (surface), and bulk materials. The smaller the unit cell size is, the faster the calculation. Modeling chemisorption at higher coverages is therefore typically easier than modeling the low-coverage situations. The small unit cell, however, can lead to inaccuracies due to enforcing periodic changes at short length scales. This can lead to high artificial local stresses that cannot be dissipated over the very small unit cell. This will likely lead to problems at higher coverages where there are adsorbate-adsorbate interactions between unit cells. The low-coverage models (larger unit cells) are typically much slower to calculate due to the increased number of atoms. These systems can be reduced since fewer k-points are needed to describe electronic structure and still maintain accuracy.

The advantage of the periodic approach is that a much more complete electronic structure can be accounted for. The disadvantages are that only systems with highly periodic structures can currently be examined within reasonable computational times. Industrial systems, however, in reality are often ill-defined supported particles, which may be more like clusters than pristine crystal surfaces.

Periodic methods are relatively new and therefore they lag behind the current cluster MO and DFT methods in terms of application. Calculating frequencies and performing transition state searches are not fully automated. This limitation will change as advances and newer algorithms are rapidly occurring. Modeling chemical reactions requires larger unit cell sizes, which can dramatically increase CPU requirements. Current state-of-the-art methods include linear response methods to model frequencies and potential energy transition state search strategies.

Periodic calculations can roughly be divided into atom-centered approaches (Gaussians, Slater-type orbitals) and plane-wave methods. The atom-centered approach had the strongest initial following due to its rigorous treatment. However, accurate plane-wave algorithms are currently available, e.g., CASSTEP (Cambridge University and MSI), VASP (Technical University of Vienna), DACAPO (Center for Atom-Scale Materials Physics, Technical University of Denmark and University of Aarhus), and CAMP-Atami (Japanese CAMP consortium)—thus tremendously reducing the CPU times required. Many of the current methods use ultrasoft pseudopotentials that cut down computational times by a factor of two or three (Vanderbilt 1990). This development begins to make it possible to model the dynamics of surfaces at a reliably accurate quantum-mechanical level.

Summary

Cluster, embedded-cluster and periodic methods have all played important roles in understanding model chemistries. The method of choice for a particular problem clearly requires a careful consideration of the advantages and disadvantages associated with each and their impact on the resulting outcome. In the future, all three will likely be invaluable in modeling both heterogeneous as well as homogenous catalytic systems.

Molecular Simulation

The prediction of structural properties of oxides, zeolites, metals, and metal sulfides, as well as weak physisorption and transport properties for these materials, is characteristically quite different than the prediction of reactivity. Structural properties can be modeled using either molecular mechanics, simulated annealing, equilibrium Monte Carlo and even molecular dynamics methods (van Santen and Kramer 1995; Frenkel and Smit 1996; Catlow 1997; Bell, Maginn and Theodorou 1997).

The bases of these algorithms were discussed in Chapter 3, while the methods used in catalysis are outlined in Table 8.2. All of these approaches rely on accurate potentials that describe bond lengths, bond angles, torsion angles, and hydrogen bonding interactions, as well as other relevant terms. The potentials have been derived empirically for the most part, but increasingly they are developed from first-principles methods. The accuracy and applicability of the approach are highly dependent on the quality of the potentials.

Table 8.2
Summary of the Basic Molecular Simulation Methods Used in Modeling Homogeneous and Heterogeneous Catalysis

Structure and Sorption	Dynamics and Kinetics
Molecular Mechanics	Molecular Dynamics
Simulated Annealing	Kinetic (Dynamic) Monte Carlo
Grand Canonical Monte Carlo	
Molecular Docking	

There have been considerable efforts in molecular simulations of zeolites due to their tremendous importance to catalytic chemistry and separation processes. Zeolites are ideal in that their crystal structures are well established and the potentials for Si, O and Al are fairly reliable. Modeling zeolite structure is therefore now fairly well established and accepted. Structures of various new materials have been identified by comparing measured X-ray powder patterns with simulated powder patterns of structures identified from simulated annealing.

The application of these tools to metals, metal oxides, and metal sulfides has been more difficult because of the difficulty in identifying accurate interatomic potentials. Simulating metals alone is an active area of research for the materials community. Fairly accurate potentials now exist for many of the metals and have been employed in the molecular dynamics codes for metal deposition. The difficulty comes when one attempts to add in organic adsorbates. There are very few studies on simulation of adsorbates on transition metal surfaces. This is due to the difficulty in establishing accurate organic/inorganic potentials. The development of accurate transition metal oxide potentials is growing. These systems, however, are quite difficult to model due to different oxidation states that can exist for a metal throughout the simulation.

In addition to structural determination, molecular simulation methods have been used to predict sorption energies, docking, and even dynamic properties such as diffusion of organics through well-defined porous networks. There have been considerable efforts aimed at examining separation of gas mixtures using molecular mechanics and dynamics methods. MD and equilibrium Monte Carlo approaches have also been successfully used to examine the structure of metal particles.

The ability to simulate the dynamics and kinetics on surfaces and through porous networks is critical toward understanding the reactivity of different materials. This area is only just beginning to emerge. Molecular dynamics, kinetic Monte Carlo, transition state theory, and *ab initio* MD all have been used to begin modeling kinetic and dynamic features such as diffusion in zeolites, surface reactivity and microkinetics (van Santen and Kramer 1995; Catlow 1997; Bell, Maginn and Theodorou 1997; Kang and Weinberg 1995; Hansen and Neurock 2000; van Santen and Neurock 2000; Hansen and Neurock 1999). These methods appear to be quite fruitful in enabling calculations at greater time and length scales.

PRESENT APPLICATIONS

Homogeneous Catalysis

Most of the work currently being carried out in industry in the area of homogeneous catalysis is aimed at elucidating reaction mechanisms and aiding the design of new catalytically active organometallic complexes. Modeling of metallocene and other homogeneous catalysts has been one of the most productive and successful activities of molecular modeling in the companies this WTEC panel visited. By way of example, we cite some of the work performed at Union Carbide below. The efforts at many other sites are similar and described in the site reports in Appendix B.

Most of this work is devoted to developing metallocene or metallocene replacement catalysts (Hackmann and Rieger 1997; Margl, Deng and Ziegler 1999; Woo et al. 1999). Metallocenes, as described above, are a general class of organometallic complexes with a single metal atom (single-site) sandwiched between one or two cyclopentadienyl-centered ligands. A number of papers have been published which examine the elementary reaction steps, such as initiation, propagation, chain coupling, and chain termination, and the effect of changing the central metal atom or the ligands attached. Metallocene-catalyzed olefin polymerization is thought to follow the general Cossee-Arlman mechanism (Hackmann and Rieger 1997; Margl, Deng and Ziegler 1999; Woo et al. 1999) (Fig. 8.5). Although the exact steps are still debated in the literature, there is enough known about the structure at the active site and plausible reaction paths that would enable one to think about the design of new complexes. However, much of the effort focused on designing new complexes remains proprietary.

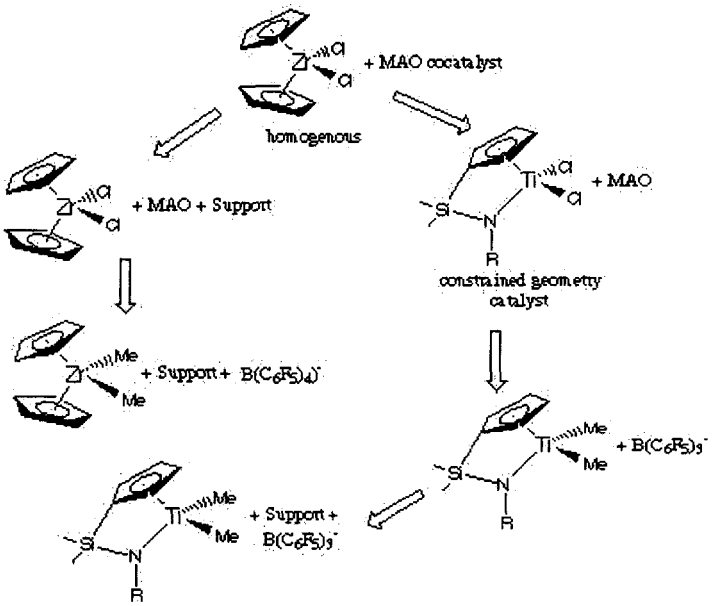


Fig. 8.5. The elementary steps in the classic Cossee-Arlman mechanism for olefin polymerization over single-site catalysts.

A few systems are presented here as examples, illustrating what is currently possible and where the field is heading. There have been a number of elegant studies published in the literature that probe the mechanism and begin to offer ideas into the design of new complexes.

One illustrative example is from the work of Professor Tom Ziegler and his colleagues at the University of Calgary (Canada) (Hackmann and Rieger 1997; Margl, Deng and Ziegler 1999). This group has been quite successful in modeling homogeneous single-site olefin polymerization catalysts. They have carried out a comprehensive analysis of the reaction chemistry for ethylene polymerization over d^0 and d^0f^n metal complexes. As a result, they have been able to identify a full sequence of elementary steps that include the structure as well as the energy of all reactants, intermediates, transition state complexes and products. Both

static as well as *ab initio* (DFT) molecular dynamics were performed on the relevant elementary reaction steps in order to isolate the transition states and identify their corresponding activation barriers.

For example, the results in Fig. 8.6 depict a transition-state structure, the potential energy surface and the heats and entropies of activation from classical ADF quantum chemical calculations, *ab initio* molecular dynamics simulations (DFT-PAW), and experimental results for the migratory insertion of methyl group into a M-CO bond in $[M(CO)_2I_3(CH_3)]^-$ where M=Ir and Rh. This is one of the elementary steps in the carbonylation of methanol to acetic acid (Cheong, Schmid and Ziegler 2000). The results from ADF cluster calculations and *ab initio* MD simulations were found to be quite similar. The *ab initio* MD approach follows a constrained free energy path between reactant and product state and offers a detailed analysis of the nature of the path as well as the overall activation barriers and free energy. While the classical cluster DFT calculations and the *ab initio* MD method provide similar results, as shown here, there are a number of other cases where the *ab initio* MD will provide the only reasonable results.

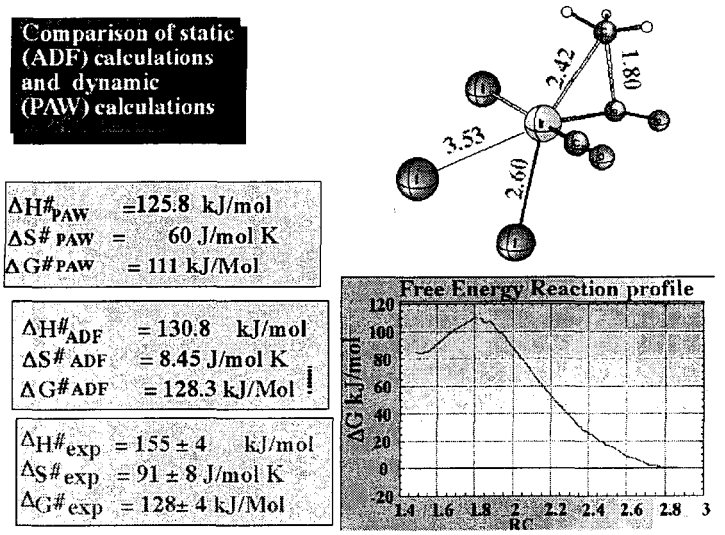


Fig. 8.6. Activation barrier predictions for CHx insertion for olefin polymerization. The comparison between traditional DFT isolated search strategies (using ADF) and *ab initio* DFT dynamic simulations (adapted from Web pages of the Ziegler group, http://www.cobalt.chem.ucalgary.ca/group/).

This group has also examined in detail the mechanism of olefin polymerization, which involves the adsorption of olefin at the metal center, olefin insertion into the growing polymer chain, termination of the growing hydrocarbon chain, and ejection of the product from the metal. A full suite of different metals and ligands were systematically examined in order to optimize the organometallic complex for activity (Margl, Deng and Ziegler 1999; Woo et al. 1999). A representative set of d^0 and d^0f^n metals was analyzed including Sc(III), Y(III), La(III), Lu(III), Ti(IV), Zr(IV), Hf(IV), Ce(IV), Th(IV), and V(V), which follow the general form of the $\{L\}M\text{-}C_xH_y^{(0,+1,+2)}$ complex. The ligands were varied in order to probe optimal activity. The ligand set included L=$NH\text{-}(CH)_2\text{-}NH^{2-}$, $N(BH_2)\text{-}(CH)_2\text{-}(BH_2)N^{2-}$, $O\text{-}(CH)_2\text{-}NH^{2-}$, Cp_2^{2-}, $NH\text{-}Si(H_2)\text{-}C_5H_4^{2-}$, $\{(oxo)(O\text{-}(CH)_3\text{-}O)\}^{3-}$, $(NH_2)_2^{2-}$, $(OH)_2^{2-}$, $(CH_3)_2^{2-}$, $NH\text{-}(CH_2)_3\text{-}NH^{2-}$, and $O\text{-}(CH_2)_3\text{-}O^{2-}$. The ligands are small and primarily probe only the electronic effects. Actual industrial systems typically use bulkier substituents that impart steric effects.

The mechanism for olefin polymerization is sketched in Fig. 8.7. Ethylene adsorption was found to favor a π-adsorption complex. The structure about the metal center adopts a trigonal planar arrangement of the ligands in order to reduce steric effects. Olefin uptake was found to correlate nicely with the charge on the metal ion coordination center and as well with the accessibility of the metal to the approaching olefin. These correlations are useful in predicting the trends for olefin adsorption. There is an optimal range of complexation energies whereby ethylene is strong enough to remain bound to the metal but weak enough to undergo chain growth reactions. Figure 8.7 can be used, for example, to begin to establish a set of possible lead candidates. Olefin complexation is important for predicting selectivity as well as the tacticity of the resulting polymer.

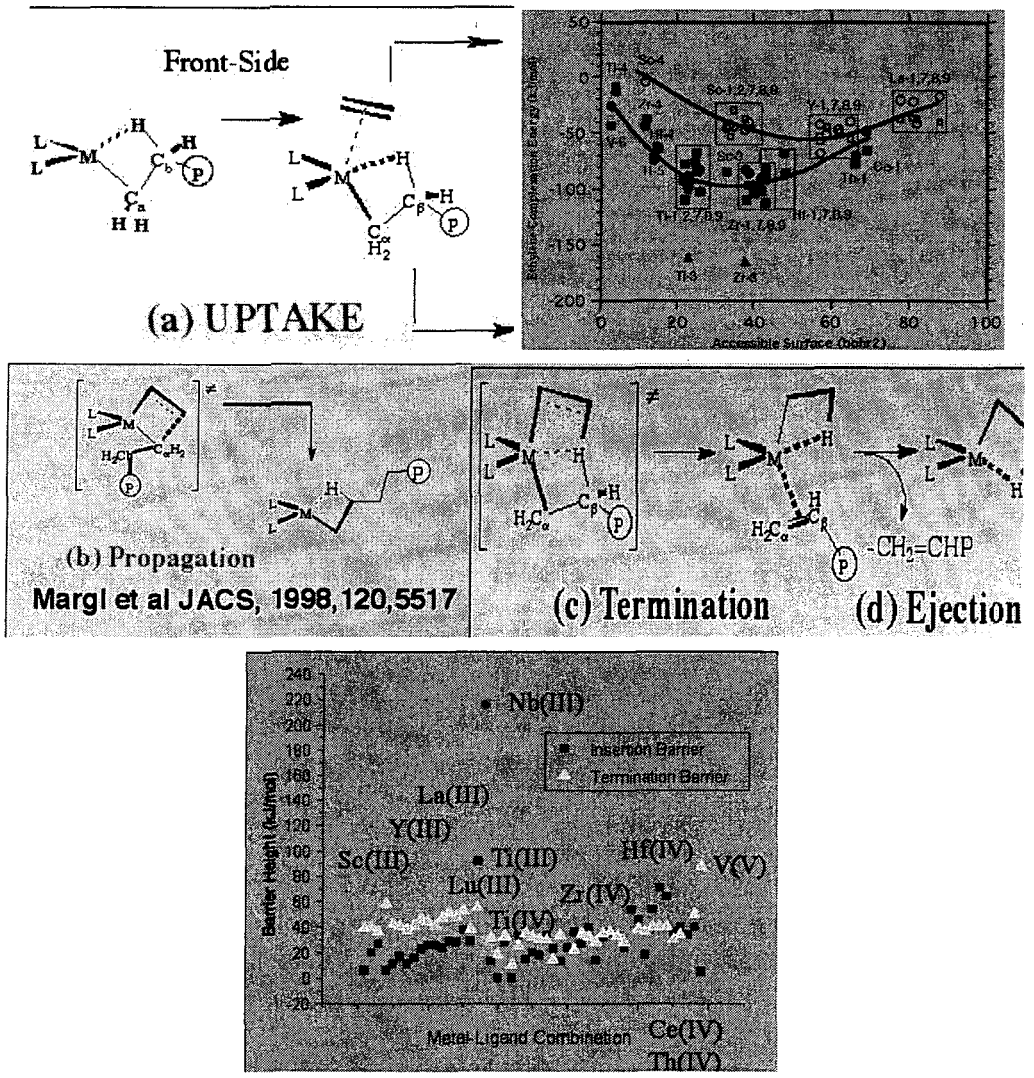


Fig. 8.7. Metal design for metallocene olefin polymerization catalysis: C_2H_4 adsorption or complexation on model ML_2 complexes (top left); effect of the metal and its oxidation state on C_2H_4 complexation (top right); chain propagation (center left); chain termination and chain ejection (center right); periodic trends of the barrier heights for olefin insertion and chain termination steps (bottom) (Margl, Deng and Ziegler 1999; Woo et al. 1999).

The optimal design of the metal center requires an intricate balance of metal-olefin and metal-intermediate bond strengths. An optimal complex will have a high activation barrier for catalytic termination and a low barrier for ethylene insertion. Comparison of chain initiation and chain termination steps for different metals and their oxidation states is shown in Figure 8.7. This work nicely illustrates what is currently possible in terms of materials optimization and design for homogeneous systems. Similar efforts are also being carried out in industry aimed at the development and design of new metallocenes.

There are now multiple industrial patents where computational chemistry has been used to screen a large number of potential ligands in an effort to establish new leads in the design of homogeneous catalytic complexes. For example, Dr. Jack Smith of Union Carbide identified active catalyst compositions (Smith 1999) which included: (1) an organometallic complex made up of a single metal atom center and a non-cycloalkadienyl anionic ligand and (2) an activation cocatalyst. A large number of potential candidates were screened using semi-empirical MOPAC calculations to calculate two relevant properties: chemical hardness and the chemical potential of the ligands. These properties were ultimately tied to the reactivity of the complex. The goal was to identify non-metallocene organometallic systems that would have similar catalytic

performance, thus expanding the range of materials and features toward the design of new catalysts. Ligands were designed in order to meet the relationship between chemical hardness and chemical potential $\eta=3.0 + 0.75\mu$ +/- 0.1 eV. The chemical potential were constrained to fall within the range of –0.5 to 3.0 eV (Smith 1999). Organometallic complexes with these ligand properties were speculated to take on similar activity to the base metal complex.

By manipulating the ligands, Smith was able to identify a range of different complexes that would lead to different polymers with different properties. The ligand takes the form of one of the general structures shown in Fig. 8.8. The groups X and X' were selected from a set of organic functional groups including CRR', SiRR', NR, PR, O, and S. The groups R and R' were selected from a set of functional groups including allyl, aryl, and heterocyclic groups. The activating co-catalyst can be a branched or cyclic oligomeric poly(hydrocarbyl)aluminum oxide, a boron alkyl, or preferably an aluminoxane. A number of MOPAC calculations were subsequently performed in order to calculate the highest and lowest occupied orbitals. This was used to specify the hardness along with chemical potential that was ultimately used to establish the likelihood for the complex to exhibit high catalytic activity. The approach was used to identify novel non-metallocene catalytic complexes.

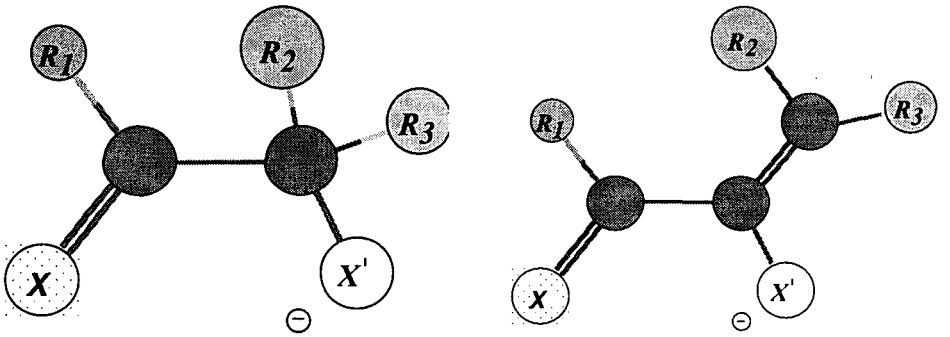


Fig. 8.8. Ligand structures used in the modeling approach of Union Carbide (Smith 1999).

Heterogeneous Catalysis

Modeling of heterogeneous catalytic systems is much more difficult than modeling homogeneous systems. Little is known about the active surface structure for most heterogeneous catalysts. In addition, the exposed surface structure can change with time as the exposed processing environment changes. Zeolites and certain other crystalline oxide materials are well-known exceptions in that they have well defined architectures.

Most of the effort in modeling heterogeneous systems has therefore focused on analysis rather than design. Academic research has been distributed over metals, metal oxides, metal sulfides and zeolites. There are also some initial efforts on more exploratory materials such as metal nitrides and carbides. Much of the early industrial effort, however, was concentrated on zeolites. The necessity to demonstrate direct relevance and have an impact on a company's efforts precluded efforts in the more ill-defined materials such as metals, oxides and sulfides. This is rapidly changing, however, as many of the more established industrial modeling groups are now also working on metals and metal oxides.

Currently, much of the work in modeling heterogeneous catalytic systems has been devoted to determining structure, predicting chemisorption, developing mechanism, and establishing structure-property relationships. Molecular modeling is currently used in conjunction with experimental X-ray diffraction, EXAFS, and XANES techniques in order to help resolve the structure. Zeolites are ideal in that they are crystalline materials with bulk properties that readily relate to their atomic structure. Most of the earliest success stories in heterogeneous catalysis relate to the combination of molecular modeling analytical characterization including X-Ray crystallography, ^{29}Si-NMR, in order to establish the structure of zeolites. Successful studies have often followed the sequence:

- Determine structure
- Predict chemisorption

- Elucidate reaction mechanisms
- Establish structure-property relationships

One example is from Union Carbide, where scientists were able to identify a new material that had known catalytic activity and ultimately received a composition-of-matter patent (Schreck 2000). Similarly, researchers at EniTecnologie S.p.A. in Italy and the University of Santa Barbara were able to establish the structure for ERS-7 (shown in Fig. 8.9). They used MSI's simulated-annealing code to find a series of lowest energy structures and an algorithm to simulate diffraction patterns in order to match the simulated and experimental powder patterns. Enichem patented the zeolite as a new material (Eniricerche 1994). Currently, there are many different techniques, including Grand Canonical Monte Carlo simulation, Monte Carlo simulation, molecular dynamics, *ab initio* MD and simulated annealing, that have been used to probe the zeolite structure, the location and distribution of aluminum atoms, and the substitution and exchange of counter ions. There are a number of very good reviews that describe these efforts in more detail (van Santen and Kramer 1995; Frenkel and Smit 1996; Catlow 1997; Bell, Maginn and Theodorou 1997).

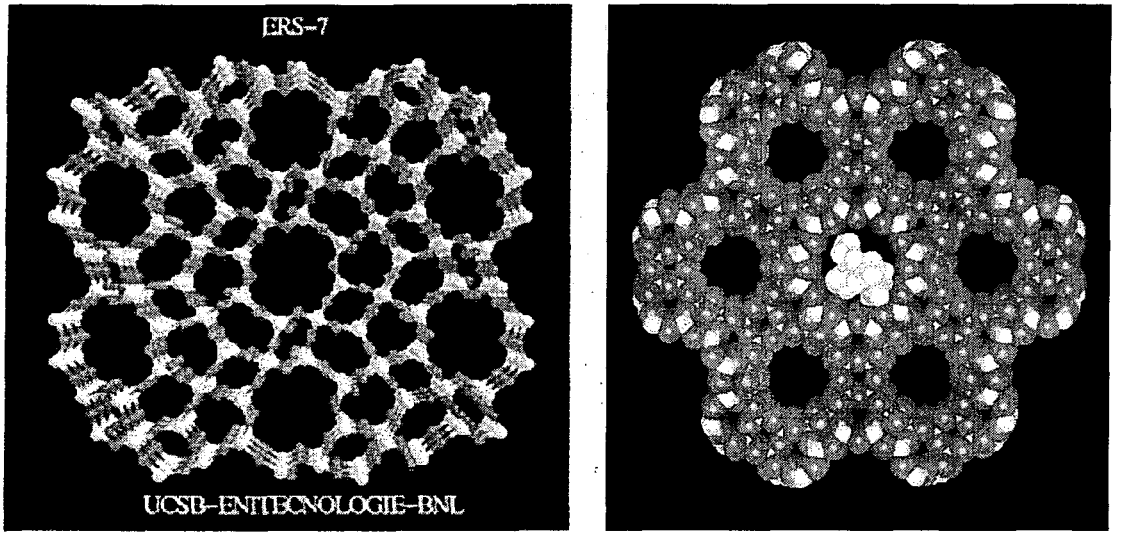


Fig. 8.9. Current simulation software is used to predict structure as well as sorption properties in zeolites: ERS-7 structure (left) (Schreck 2000) was established using simulated-annealing algorithms to match X-ray powder diffraction patterns. Site location and energies were predicted by Bell and Catlow (Bell et al. 1994; Sankar, Thomas and Catlow 2000; Sankar et al. 1998; Barrett et al. 1996) for *n*-octane sorption in DAF (right).

Zeolites have well-defined pore structures and are composed of a well-defined array of Si, O, and Al atoms, all of which can be easily modeled using established force fields. Adsorption and diffusion of organic molecules in zeolites can be examined with a reliable level of accuracy. Much of the effort in industry and even in academia is targeted toward predicting adsorption properties, which is important for both separation and catalysis.

Catlow and Bell, for example, have studied the adsorption of various alkenes and alkanes in different microporous materials. Figure 8.9 (right) shows the preferred adsorption of *n*-octane in the microporous aluminum phosphate DAF (Bell et al. 1994; Sankar, Thomas and Catlow 2000; Sankar et al. 1998; Barrett et al. 1996). DAF was designed by Davies' research labs at the Royal Institution. In order to convert methanol selectively to ethylene and propylene (Sankar, Thomas and Catlow 2000; Sankar et al. 1998; Barrett et al. 1996), Grand Canonical Monte Carlo simulations, docking, and molecular dynamics simulation have been used to establish the best sites for adsorption, the strength of adsorption, and the equilibrium between adsorbed phases. Kinetic Monte Carlo, transition-state theory and molecular dynamics have been used to explore the diffusion of organic molecules through porous zeolite architectures. The current state of the art enables modeling of adsorption and diffusion of single species (Neurock and Bell 1997; Catlow 1997; Bell, Maginn and Theodorou 1997). Work is ongoing, but scientists have yet to establish an effective route to describe multicomponent diffusion and reactions.

In the area of metal and metal oxide catalysis, much less is known about the active surface structure. Many of the published papers therefore attempt to compare modeling results with experiments on well-defined single crystal surfaces, which are examined under pristine ultrahigh vacuum conditions. There is now a growing database for the structure and chemisorption energies of various adsorbates on different metal and metal oxide surfaces that have been calculated. The results agree rather well with known experimental results. The error bars for predicting adsorption energies on these surfaces are usually in the range of 5-8 kcal/mol. This was shown earlier in the parity plot of DFT-calculated binding energies of different adsorbates on well-defined metal surfaces compared against experimental results, Fig. 8.3 (left).

In addition to determining the structure and energies for the reactants, it is also quite easy to examine the nature of adsorbed intermediates and products as well. The predicted energies can subsequently be used to establish potential energy profiles for the overall reaction schemes. The cluster, embedded cluster and periodic slab calculations all have the capability to perform transition-state searches. There has been a tremendous growth in the past five years in the application of first-principles-based methods to isolating transition states and establishing activation barriers. This, however, is still a developing area. Isolation of transition states is at best a very tedious task. The accuracy in predicting activation barriers is thought to be similar to that for adsorption energy predictions. The database, however, is not large enough yet to establish a more precise assessment of accuracy.

The ability to calculate the reaction energies for elementary steps, along with their activation barriers for different paths, makes it possible to begin to establish likely mechanisms and, in addition, to calculate the relative degree of selectivity. For example, van Santen et al. were able to follow a sequence of plausible steps along with their activation barriers for the conversion of methanol to methyl-tertiary-butyl ether. The results suggest that the favored route occurs via methanol dehydration followed by associative coupling mechanisms (Blaszkowski and van Santen 1996). The results are presented in Fig. 8.10, which shows how these steps make up the overall catalytic cycle. This path was found to be significantly more favored than another route in which the formation of a surface methoxy intermediate occurs. More recent calculations using extended periodic slab models have shown similar trends with slightly different energetics, suggesting the pore structure can influence properties of the material (Sandre, Payne and Gale 1998; Stich et al. 1999). The ultimate objective of understanding the active site and how it operates is to control catalytic chemistry.

While tremendous progress has been made in understanding the structure, modes of chemisorption, and reactivity on well-defined systems, elucidating what happens in the reactor requires a much more realistic representation of the reaction environment under operating conditions. Much of what has been established using theory over the past few decades has focused on understanding the intrinsic metal-adsorbate bonding. However, extrinsic factors such as defect sites, higher surface coverage, promoters, poisons, and solvent effects can be just as important in dictating catalytic reactivity. A number of elegant studies are beginning to appear which begin to treat these extrinsic factors.

In the area of metals, Nørskov, Hammer, Chorkendorff, and Besenbacher have combined theory and experiment in order to understand the activity of different metals for N_2 dissociation, which is a critical step in the synthesis of ammonia (Dahl et al. 1999). This group from the Center for Atomic-Scale Materials Physics, Denmark (see site report in Appendix B) has worked quite closely with the Haldor Topsøe Company. They have shown that much of the low-temperature activity for N_2 activation occurs at defect sites on the metal. A large discrepancy existed between experimental results and the theoretical predictions for N_2 activation over ideal Ru(0001) surfaces. The barrier predicted from theory was found to be over 90 kJ/mol higher than that derived from experiment. To resolve these differences, the theoreticians calculated N_2 activation at step edge defects. The barrier was found to be significantly lower and in very good agreement with that measured experimentally. To help confirm that the "ideal" Ru(0001) surface contained reactive step edges, Chorkendorff sputtered gold onto the surface to shut down any of these defect sites (Dahl et al. 1999). The measured activity was in much better agreement with that found from the DFT predictions on the terrace planes of Ru(0001). The results are summarized here in Fig. 8.11. This example nicely illustrates the complementary role that theory and simulation are playing with regard to experiment and establishing what happens on surfaces that begin to move away from the ideal single crystals.

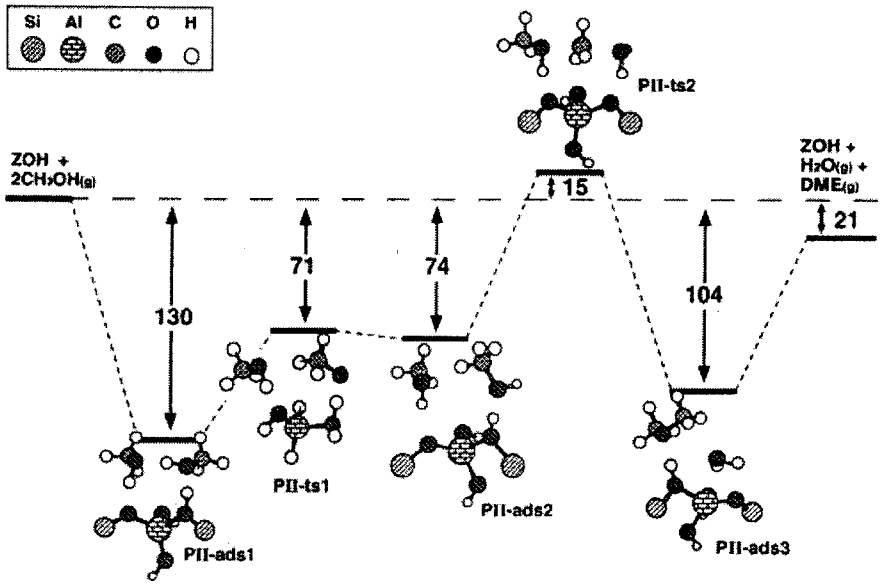


Fig. 8.10. Overall catalytic cycle for methanol to methyl-tertiary-butyl-ether catalyzed by acid catalysis in a zeolite (Blaszkowski and van Santen 1996).

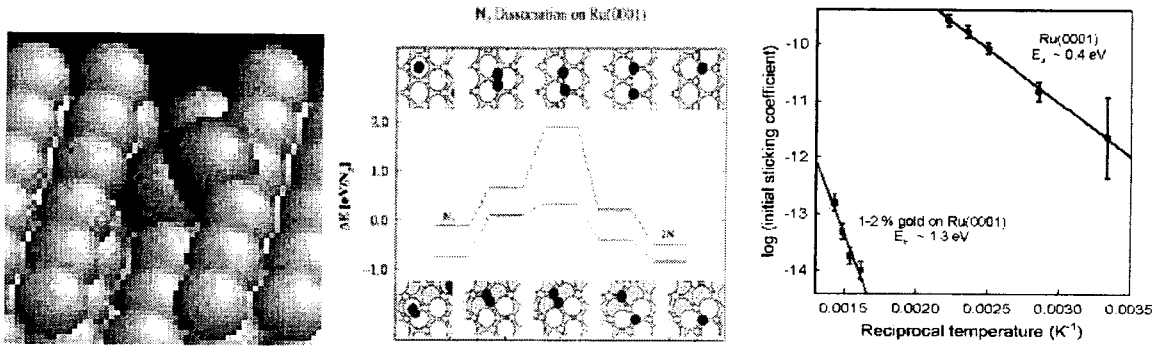


Fig. 8.11. Activation of N_2 on the Ru(0001) terraces and step edges: the dissociation of N_2 at the Ru(0001) step edge (left); DFT-calculated barriers on the terrace and at the step edge (center); and predicted activity for N_2 activation on the terrace and at the step edge (right) (Dahl et al. 1999).

The nature of the chemical environment can critically affect the stable surface structure. In the area of oxides, Scheffler et al. (Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin) (Wang, Chaka and Scheffler 2000) have pioneered an approach that uses first-principles DFT calculations to determine the free energies of the surface structure in various different environments. By calculating the chemical potentials at the operating conditions, they are able to establish the most stable surfaces that can exist and establish the likelihood of chemisorbed species such as hydroxyl groups. A phase diagram, presented in Fig. 8.12, follows the favorable surfaces for Al_2O_3 as the oxygen partial pressure is increased.

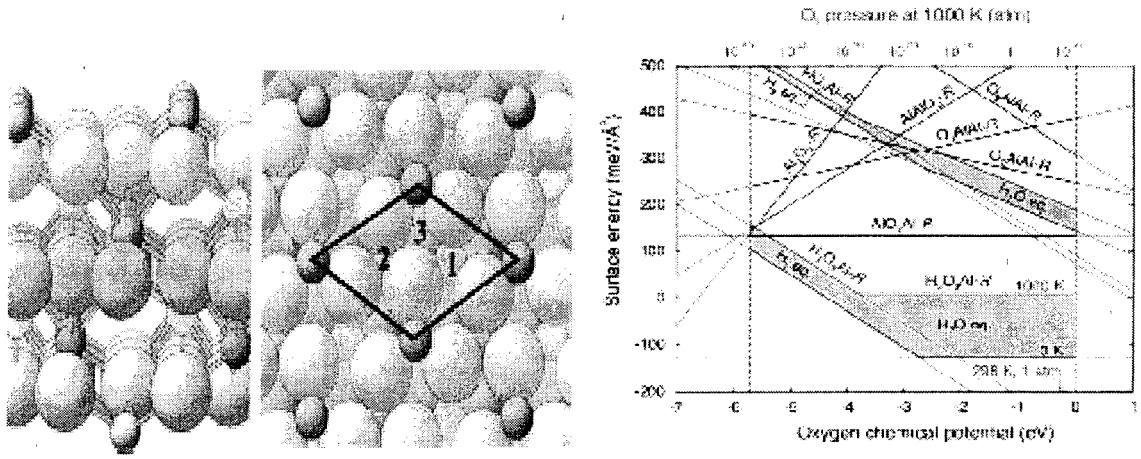


Fig. 8.12. The effect of gas-phase oxygen partial pressure on the stability of the surface structure: Al_2O_3 surface oxide (left); a phase diagram for the most stable phases (right).

There are a number of other elegant studies in the literature that also target the extrinsic factors that focus on the reaction environment as was discussed earlier. It is clear that future efforts will require the ability to understand and predict both the intrinsic and extrinsic factors that control catalytic activity.

The simulation of chemistry on metals and metal oxide surfaces is more difficult than that for zeolites. Both MD and Metropolis Monte Carlo tools, which are useful in simulating sorption and diffusion in siliceous materials, have been underdeveloped for transition-metal catalysis. The difficulty is that accurate potentials for transition metals and organics together have yet to be developed. There has, however, been progress in the use of kinetic or dynamic Monte Carlo methods (Kang and Weinberg 1995; Neurock and Hansen 1998). Hansen and Neurock (2000; 1998; 1999) have shown that first-principles DFT calculations can be used to establish the kinetic parameters, chemisorption constants, heats of reaction, and the interaction energies between different surface adsorbates in order to develop a comprehensive database. This database provides the necessary input to a kinetic Monte Carlo simulation, which can track a complex sequence of elementary steps that comprise the catalytic cycle as a function of time, or processing conditions. The simulation follows the spatiotemporal evolution of the surface structure and can thereby begin to account for both the intrinsic as well as the extrinsic features in the mechanism. The approach has been used to simulate the hydrogenation of ethylene and the synthesis of vinyl acetate over various bimetallic surfaces. A snapshot taken at a specific instant in time for the synthesis of VAM on a PdAu alloyed surface is shown in Fig. 8.13 (Hansen 2000). The picture nicely illustrates that there is very little ethylene on the surface. The surface is effectively shut down to ethylene adsorption due to the high coverage of acetate and surface oxygen that dominate the surface. The addition of gold helped to boost selectivity from 85% to 93%. The results have been used to provide suggestions to DuPont (see site report) on ways in which to improve both activity and selectivity by tailoring bimetallics.

A second application of molecular modeling toward the design of heterogeneous catalysis is that by the combined theory/experimental/industrial team in Denmark. The STM group of Besenbacher et al. (1998) found that they could generate unique NiAu surface alloys on an idealized single-crystal metal whereby the Au was found to be well dispersed. Theoretical calculations indicated that it was possible to generate such an alloy (which previously was not considered since the alloy was not stable in the bulk) and that the alloy should enhance the selectivity for steam reforming. One of the main limitations of steam reforming is that it is typically carried out over Ni metal surfaces. Nickel is known to rapidly deactivate due to coke build-up on the catalyst surface. The calculations indicated that by alloying the surface with Au, carbon formation, which leads to deactivation, would no longer be a problem.

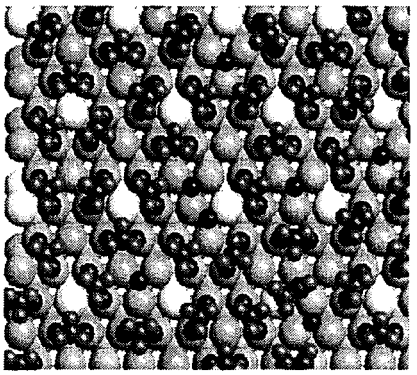
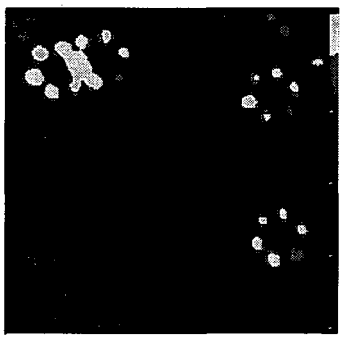


Fig. 8.13. Snapshot of the DFT-based kinetic MC simulation for the synthesis of vinyl acetate from ethylene and acetic acid over PdAu alloy. The addition of gold helps considerably to enhance selectivity and marginally to increase activity (Dahl et al. 1999).

The results of the academic experimental and theoretical groups convinced the researchers at Haldor Topsøe to try to stabilize such a surface alloy. Haldor Topsøe (see its site report) was able to create a NiAu alloyed surface which proved to be stable over prolonged operating conditions. The results are presented in Fig. 8.14. STM pictures of the Ni with well-dispersed Au under UHV conditions are shown at the top of Fig. 8.14. The simulation results, depicted in the lower left, illustrate an increase of the barrier for the unselective decomposition of methane as gold is alloyed into the surface. The final results, at the bottom, show the improvement in eliminating coke deposition problems.



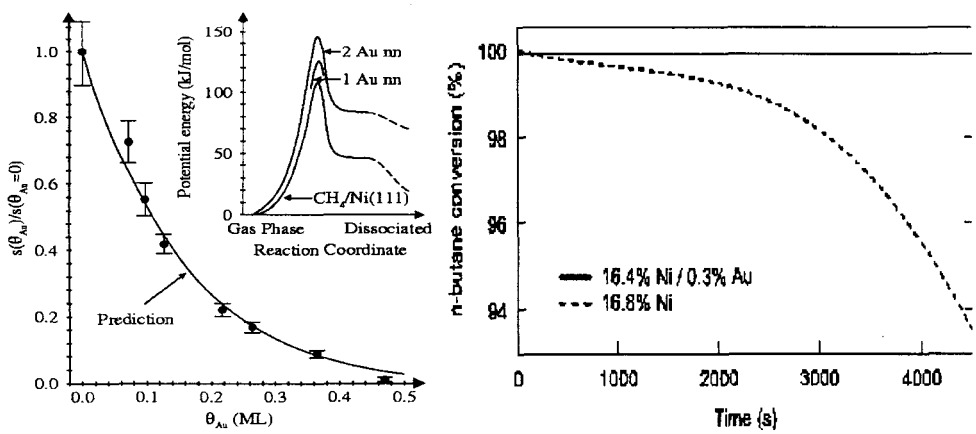


Fig. 8.14. The coupled application of surface science experiments, DFT quantum chemical calculations, and experimental synthesis in the design of a new stabilized NiAu steam-reforming catalyst (Besenbacher et al. 1998): A unique NiAu alloy is stabilized at the surface as determined from STM experiments (top). The inset graph on the lower left reveals the increase in barrier for methane activation in the presence of different NiAu alloys. The experiment shows improvement in n-butane sustainability for steam reforming (right).

PANEL FINDINGS

Overview of Catalysis in Site Reports

Tables 8.3a through 8.3c summarize industrial and government modeling in homogeneous and heterogeneous catalysis. Again, much of the work that has been performed has been focused on homogeneous and zeolitic systems.

Table 8.3a
European-based Activities In Molecularly Based Catalysis Modeling, Industrial and Government Sites

Site	Activity
Air Liquide, France	Heterogeneous catalysis, novel ligand design, zeolite design for separation
BASF AG, Germany	QM/MM, CP-MD and Classical QM applications to homogeneous, heterogeneous, and enzymatic catalysis
Bayer AG, Germany	Homogeneous catalysis (Lohrenz 1998)
Advantica Technologies Ltd. (formerly named BG Technology Ltd.), UK	Natural gas hydrates, gas separation, predicting x-ray diffraction patterns, predicting defect structures from dopant incorporation
Daresbury Laboratory, UK	Quantum Simulation in Industry (QUASI) project: enzyme catalysis applications with BASF as the lead partner; surface catalysis applications with ICI as the lead partner; and zeolite catalysis applications with Norsk Hydro as the lead partner
Degussa-Hüls AG (now Degussa), Germany	Work in homogeneous catalysis, but strong interest in heterogeneous catalysis
DSM Research, The Netherlands	Homogeneous catalysis, CP-MD metallocene-catalyzed ethylene polymerization (Meier et al. 1994; Aagaard, Meier and Buda 1998; Meier 1999), structure-reactivity relationships. Examples: Zirconocene ethylene polymerization, Ru-catalyzed olefin metathesis.
EniChem S.p.A., Italy	Metallocenes and polymer properties are a central focus (Fusco et al. 1996), interested in heterogeneous catalysis. Modelers' proposal for a modified homogenous catalyst has resulted in two patents. Molecular modeling was also used to improve another existing catalyst, which also resulted in patents. Close relationship between modelers and experimentalists. Examples: Ti, Zr Ziegler-Natta olefin polymerization.
EniTecnologie, Italy	Catalysis modeling fully integrated into multidisciplinary team. Qualitative MM/MD rather than quantitative or QM approaches. Heterogeneous—molecular mechanics analysis of catalytic shape selectivity in the syntheses of cumene, isobutene, and 2,6-dimethyl naphthalene. Modeling resulted in patents. Short-term efforts on fast catalyst screening. Zeolite structure and synthesis, including ERS-7.
GdR 12090 project, French-based multinational activity	CNRS-sponsored consortium of industry, universities, and government labs whose purpose is to develop catalysis modeling tools for the petroleum industry. Participants include CNRS, ISMRA, Institut Français du Pétrole, TOTALFINA, U. Henri Poincaré (Nancy), Université Pierre et Marie Curie (Paris), Université de Poitiers, T. U. Eindhoven, and Universitat Wien.
Haldor Topsøe, Denmark	Heterogeneous catalysis: ammonia synthesis, steam reforming, methanol synthesis. Strong collaborative efforts with Nørskøv, Chorkendorff, Besenbacher at CAMP-DTU.
Institut Français du Pétrole (IFP), France	Catalytic materials modeling, hydrodesulfurization, adsorption in microporous materials, reaction mechanism (effects of sterics and poisons), poisoning of hydrogenation catalysts
Totalfina (now TotalFinaElf), France and Belgium	Hydrodesulfurization, alkane transformations (cracking, isomerization), catalyst characterization (acidity, NMR, IR probes), interests in homogeneous catalysis
Solvay Research, Belgium	Ziegler-Natta catalysts and metal clusters

Table 8.3b
Japanese-based Activities in Molecularly Based Catalysis Modeling, Industrial and Government Sites

Site	Activity
Asahi Chemical Industry Co., Ltd.,	Homogeneous catalysis: develop trends and reliable predictions. One of the best success stories is that for the design of a single-site Zr-metallocene olefin polymerization catalyst. Computational chemistry led the development of a new catalyst and was used to help patent the material. Heterogeneous catalysts: structural analysis by force field, DFT and experimental studies.
Japan Polyolefins, Ltd.	Metallocene catalysis
Joint Research Center for Atom Technology (JRCAT)	Surface processes and chemical reactions with a focus on Ziegler-Natta catalysts for ethylene polymerization. Mechanism and reaction pathway development. Collaboration with Parrinello (Stuttgart).
Mitsubishi Chemical, Yokohama Research Center	5% of the patents from Yokohama make use of computational chemistry or molecular modeling. Molecular design of homogeneous catalysts. The design of metallocene catalysts has had an important impact. Catalysis scientists are now performing many of their own calculations. Heterogeneous catalysis.
Mitsui Chemical	Catalyst design
NEC Tsukuba Laboratories	Catalytic destruction of dioxin by dissociative adsorption on silicon. Collaborate to develop CAMP-Atami first-principles molecular dynamics code and lead CAMP project. Develop heavy-atom pseudopotentials.
Nippon Zeon Co., Ltd.	Catalysts for polymer synthesis
Sumitomo Chemical Co., Ltd.	Ziegler-type catalysts
Toshiba Corporation	Exploring dendrimer catalysts
UBE Industries, Ltd.	Polymerization catalysts

Table 8.3c
U.S.-based Activities in Molecularly Based Catalysis Modeling, Industrial And Government Sites

Site	Activity
3M, Inc.	Catalysis modeling is in a mostly experimental group, linked also to ASPEN process modeling activities for cost evaluation. Design of catalyst structures that will produce desired product structures.
Air Products and Chemicals, Inc.	Homogeneous catalysis, heterogeneous catalysis applications to amine synthesis
Amoco (now BP)	Homogeneous catalysis
The Dow Chemical Company	Developed a strong patent position in homogeneous catalysis (LaPointe et al. n.d.; McAdon et al. n.d.; Nickias et al. n.d.) through interactions of the modeling and experimental groups. (note that Dow has acquired Union Carbide)
E. I. du Pont de Nemours and Company	Both research and development for heterogeneous catalysis; heterogeneous catalysis mostly by academic collaboration
Marathon Oil Company	Quantum chemistry and QSAR for hydrogenation catalysis (activity ended December 2000)
Phillips Petroleum (now Chevron Phillips Chemical Company)	Fundamental work on homogeneous catalysis performed with the Minnesota Supercomputer Center, the University of Oklahoma, and Emory University under the NIST ATP program
Shell Oil Company	Homogeneous catalysis
Union Carbide (now part of Dow)	Homogeneous and heterogeneous catalysis described in Chapter 8

General Trends—Homogeneous Catalysis

Computational chemistry continues to play an important role in homogeneous catalysis R&D efforts. Much of the initial effort was aimed at understanding the chemistry and elucidating mechanism. Over the past five years, theory has taken on a more active role in the design and optimization of the metals and ligands for homogeneous catalytic systems. There have been many success stories whereby computational chemistry was used to strengthen intellectual property (IP) in homogeneous catalysis—it can be used both in developing and defending IP. This use of modeling to help clarify IP rights in homogenous catalysis is clearly a growing trend.

In most of these companies, computational chemistry has been accepted for modeling homogeneous catalytic systems because it has helped them or their competitors establish a solid foundation of credibility and reliability in predicting known thermochemical information and offering acceptable new leads. The success rate continues to improve and increase the internal support for modeling. This trend also has helped to increase the number of new hires of modelers at some of the companies visited. The success of homogeneous efforts at Union Carbide, for example, has led to hiring personnel to examine heterogeneous systems as well.

The modeling currently carried out in industry involves the application of molecular mechanics as well as various different quantum chemical methods, including semi-empirical, HF-CI and DFT approaches. The major goals are typically (1) to isolate the active sites, (2) to calculate binding energies and activation barriers, and (3) to explore the effects of changing the metal ion or the electronic and steric nature of the ligands toward the optimal design of the organometallic complex. In addition to performing calculations, many of the leading computational chemists are also in charge of training and working with other chemists and chemical engineers. This has helped to gain more general acceptance and extend the use of modeling to a broader community. The full-time modelers have had to write useful scripts or other general-purpose code that enables more efficient ways to guide the users and to handle the large amounts of data. As high-throughput-screening has become more important, it is likely that computational chemistry will also play a very important role in library, screening and even informatics developments, as it has in drug discovery.

In many of the companies visited, the number of projects clearly outweighed the number of personnel. In addition, the short time scales given to come up with solutions has made it very difficult for many of the scientists to carry out code development. Nearly all of the code development has been done via partnering with academia or other external groups. For example, many of the applications to homogeneous catalytic systems require the description of bulky ligand systems along with core transition metal sites. In addition, many of these reactions are run in the presence of a solvent. Industry has had to rely on external collaborations or software vendors to develop the codes necessary to carry this out. This has involved the development of ONIOM-type models, molecular mechanics modules, and solvent models. In addition to method development, a few of the companies had partnered with external groups in order to devise sets of pseudopotentials for specific systems of interest.

General Trends—Heterogeneous Catalysis

It was clear that the WTEC panel's hosts at many of the companies thought that modeling heterogeneous systems is critically important for the future. Most commercial processes use heterogeneous rather than homogeneous catalysts. This is expected to increase significantly over the next decade as homogenous systems are replaced by heterogeneous systems as we move toward "greener" chemical processes. However, much less is known about heterogeneous catalysts, due to difficulty in resolving surface structure and how it changes with processing.

Despite these difficulties, most companies would clearly like to see progress in modeling heterogeneous catalytic systems. Because of the difficulties associated with modeling these systems, many of the companies have yet to fully commit to these efforts and could not justify hiring someone specifically for modeling heterogeneous systems. Some of the companies have asked their homogeneous modelers to spend a fraction of their time looking at heterogeneous systems.

However, most of the companies interviewed have formed external collaborations with either academia or national labs in order to support more modest efforts. This appears to be universal throughout the United States and Europe. The Japanese have formed still larger external collaborations, often aided by the government, among academic institutions, government institutes, and other companies. By way of example, in the United States, Air Products, BOC, Dow, DuPont, Eastman Chemical, Ford, and Monsanto have all recently supported external collaborations in modeling heterogeneous catalytic systems. In Europe, Rhône Poulenc (Rhodia), Shell, Totalfina, Solvay, Akzo, Advantica Technologies Ltd. (formerly BG Technology), and Haldor Topsøe have all had maintained strong ties via external collaboration with industry and government labs. In addition, some of the companies were actively involved in the MSI Catalysis Consortium project. While external collaboration was common, a few companies the WTEC team visited had no effort in heterogeneous catalysis. The largest internal efforts for modeling heterogeneous catalysis appeared to be concentrated at Ford, Union Carbide, and BASF.

Much of the work on heterogeneous systems has focused on structural characterization and on elucidating reaction mechanism. Most of this work has targeted zeolites, which have well-defined pore structures and which are thus ideal for modeling, as discussed above.

Current Industrial Trends

The strength of modeling in industry appears to be closely tied with the company's R&D focus. Catalysis modeling appears to mirror experimental R&D programs. For companies where fundamental research is supported, modeling efforts are strong. Those companies that have decreased fundamental research activities have also decreased their efforts in molecular modeling. Most of this is simply tied to profits and earnings. Eastman Chemical Company, for example, announced a large reduction in R&D staff in October 1999. This cut appeared to be universal, thus reducing molecular modeling as well as other research programs. This is not a major surprise. A breakdown across industry is given below:

R&D in the petroleum and gas industry, including modeling, has diminished dramatically over the past five years. This has been coupled with a steady reduction in staff during this period. Much of what is currently practiced can be considered technology transfer rather than research and development.

In chemical manufacturing and processing industries, the level of effort varies considerably among companies. A number of companies, such as Dow, Air Products, and Enichem, have witnessed strong sustainable programs. Dow, Ford and Union Carbide have all grown over the past few years by hiring to strengthen their efforts in both homogeneous and heterogeneous modeling. Union Carbide estimates that it needs at least one modeler per every five or six experimentalists to support research efforts effectively. On the flip side, several other companies, including Eastman, Akzo, and Shell/KSLA Chemical, have cancelled their molecular modeling efforts. These reductions, however, were all part of major efforts within these companies to reduce longer term R&D efforts (both modeling as well as experimental research).

There are also several companies that are dramatically altering their focus. DuPont, for example, is shifting from traditionally chemical and material sciences into life sciences. R&D catalysis efforts have shifted in order to accommodate this new push into life science and biomimetic materials. Modeling efforts are also beginning to reflect these changes as well. Many other chemical companies are also beginning to look more closely at life sciences, but few have demonstrated the same level of commitment as DuPont. Current perturbations in the chemicals community due to mergers and takeovers will also impact molecular modeling.

It is worth noting the great degree of progress in modeling that has been made by non-publicly held companies. Haldor Topsøe and Enichem have been able to capitalize on modeling for heterogeneous catalytic systems. To explain their successes, both of these companies have cited their ability to invest in modeling catalytic systems with a longer-term vision in mind. They have been able to examine some of the more difficult problems without worrying about how financial analysts would view these efforts and whether it might hurt the company's stock. This has helped them to make continuous progress on more difficult systems.

To summarize, the level of industrial commitment to molecular modeling closely mirrors what is seen in experimental R&D and is strongly tied to business decisions on company directions and research expenditures. Unfortunately, the short-term and often fragmented goals of many publicly held companies make it difficult for them to mount longer-term efforts aimed at solving the more difficult but critically important issues. This is one of the reasons why there is much less effort in heterogeneous catalysis. Both European and Japanese governments have begun to recognize the high potential for break-through ideas and technology by sponsoring industry/academia/national lab partnerships at fairly high funding levels. The Japanese Joint Research Center for Atom Technology (JRCAT) and the European GdR 12090 project are just two of examples of this.

Comparison of International Academic Efforts and Strengths

A detailed comparison among international efforts in the area of modeling catalysis is difficult in that there have been a number of very important achievements in Europe, Asia, and North America. Therefore, ranking international groups is very challenging. A close inspection and comparison of the research across Europe, Japan and the United States, however, suggests some very interesting trends. These trends can, to some extent, be explained by the backgrounds of the respective research groups involved.

The Europeans have been leaders both in developing *ab initio* solid-state algorithms and their application to the analysis of catalytic chemistry in metals, metal oxides and metal sulfides. Many of the periodic plane wave codes used today were developed in European physics departments, including CASTEP (M. Payne, Cambridge), DACAPO (B. Hammer and J. Nørskov, Denmark), VASP (G. Kresse and J. Hafner, Vienna), and Wien97 (P. Blaha and K. Schwarz, Vienna). The core strengths in modeling catalytic chemistry on surfaces have come from both the chemistry and physics communities in Europe. The European chemical engineering programs have traditionally been much more classically oriented and have emphasized traditional chemical engineering areas such as unit operations, fluid mechanics, mass and heat transfer, and reaction engineering. We did not, however, come across extensive work on the application of *ab initio* methods from the European chemical engineering programs.

Europe has also been a leader in developing interatomic potentials for modeling oxide materials. Most of this work has come from the chemistry community: R. Catlow and J. Gale (United Kingdom), B. Smit and R. van Santen (Netherlands), and J. Sauer (Germany), just to name a few. The European work in modeling homogeneous catalysis also has been quite strong.

European governments have begun to recognize their strengths in modeling heterogeneous and homogeneous catalysis. They are offering more multidisciplinary and inter-European monies in order to foster the development of quantum mechanical methods applied to catalytic systems. Examples are the United Kingdom's CP consortium and the GdR 12090 project from France, which joins scientists from France, Vienna, The Netherlands, and industry (Totalfina, Institut Français du Pétrole) in order to solve a number of difficult heterogeneous catalytic systems in petroleum processing.

Researchers in the United States and Canada have played leading roles in the development of quantum chemical cluster methods, as well as their application to homogeneous catalytic reaction systems. These efforts have come predominantly from the chemistry community. The U.S. and Canada continue to excel in these areas offering fast algorithms and the ability to treat extended systems through embedding systems such as ONIOM-type approaches. The efforts for modeling heterogeneous catalysis in the United States started in chemistry, but the number of researchers in the U.S has been far fewer than those in Europe.

U.S. physicists have contributed greatly to the discovery and development of theory. Walter Kohn (physicist) shared the Nobel Prize in Chemistry with John Pople (chemist), recognizing their role in the development of density functional theory and its impact on computational chemistry. Joannopoulos (MIT), Teter (Corning), and Allen (Corning) were responsible for some of the earliest DFT planewave code development. This work was conducted in conjunction with Michael Payne from Cambridge University (UK). However, the number of U.S. academic physicists in the area of modeling heterogeneous catalysis is very small.

U.S. academic physicists and chemists have traditionally focused more on basic science than on engineering application. This is a very general observation. While the United States has a core group of chemists and

physicists who have been active in modeling heterogeneous systems, they are somewhat few in number compared with Europe. They have all made important contributions, however.

Interestingly, in the U.S. chemical engineers have begun to move into the area of *ab initio* calculations for heterogeneous catalytic systems. The area offers a mix of physics, chemistry and engineering application. It is therefore becoming a growing area of focus for U.S. academic chemical engineering departments. These efforts, however, have been targeted primarily at the application and extension of QM methods. *Ab initio* code development is still primarily in the hands of the physicists and chemists.

The United States has played a leading role in *molecular simulation* (MD, MC, docking, etc.) of heterogeneous catalysis by being one of the leaders in the development of simulation algorithms for modeling diffusion, adsorption, and for surface reaction in zeolite channels, metal surfaces, metal oxide surfaces, and metal sulfide surfaces. Much of this effort has been out of the chemical engineering community along with some efforts in chemistry and physics. The chemical engineering community has had a history of developing molecular- simulation algorithms and therefore has easily adapted these codes to catalytic material.

The Japanese have core strengths in the applications and development of first-principles methods to homogeneous catalysis. While they have significantly fewer researchers dedicated to the application of modeling to extended surfaces for catalysis, they do have an excellent team of scientists aimed at modeling the solid-state properties of various materials, primarily for the electronics industry. The Japanese have also begun to move into molecular simulations of porous materials. The number of scientists and engineers in this area, however, is still quite small.

In comparing Europe, United States and Japan, the strengths of each country are inherently tied to the academic departments that are working in the area. The U.S. physics and chemistry communities appear to be more focused on basic sciences. U.S. chemists and physicists have contributed greatly to the development of the theory, methods and codes that are being used. In particular, many of the commercial codes such as Gaussian, Jaguar, and QChem originated in U.S. chemistry departments. Many additional features, such as linear-scaling methods and the ONIOM models, are also being developed in U.S. chemistry departments.

Chemistry and physics in Europe have a balance of both applied and basic science efforts. There are a growing number of groups in chemistry and physics in Europe that work with industry in modeling complicated metal, metal oxide and metal sulfide chemistry. Chemical engineering programs in the U.S. are somewhat more fundamental than those in Europe and, in that sense, may be a bit closer in nature to some of the European chemistry programs.

The activity in experimental catalysis is an interesting example that highlights very similar insights. Many of the experimental efforts aimed at heterogeneous catalysis in Europe are housed in chemistry, while in the United States, they are housed in chemical engineering. Once again, this is a very general observation. It should be pointed out that there have been a number of outstanding contributions from U.S. chemists and European chemical engineers. It is therefore quite logical that modeling shows the same trends. While the strengths of chemical engineering in the U.S. are much more related to molecular simulation, the effort in *ab initio* applications will grow.

NEEDS AND LIMITATIONS

Needs and limitations can be broken down into algorithmic and application-specific areas. Algorithmic needs are quite general and essentially identical to the needs presented for biological, materials and other application areas. Application needs are those that are more specifically tied to homogeneous and heterogeneous catalysis. Table 8.4 summarizes the greatest needs for method development.

Algorithmic Needs

The accuracy of modeling heterogeneous and homogeneous catalysis is coupled both to the accuracy of the method and the models employed. Both areas can be critically improved, as discussed earlier, by the

development of faster, cheaper, and more accurate methods. These methods involve linear scaling, divide-and-conquer schemes, improved transition-state search algorithms, more accurate and appropriate functionals for density functional methods, mixed CI/DFT schemes (which allow for very high-level CI treatment of the core reaction site along with faster DFT methods to describe the surroundings), improved algorithms to predict properties, and accurate semi-empirical calculations for transition metal systems. There is on-going work in all of these areas world-wide.

Table 8.4
General Industrial Computational Needs

Faster, cheaper and more accurate methods	Order n-scaling, improved transition state search algorithms, better functionals for DFT, improved property prediction, accurate semi-empirical methods
Ability to model more realistic environments	Chirality, defects, promoters/poisons, solvents, electric fields supports, multimetallics, multifunctional materials
Seamless integration between time and length scales	Electronic, molecular, meso, macro, and system scales
Robust force fields and robust simulation tools for broad application to materials	Reactive FF, charge transfer, induced polarization

Accuracy, and in some cases even our ability to model certain systems, is controlled by how realistic a model of the reaction environment we can create. The reaction environment model must be large enough to capture the important features about the active site, yet small enough to be able to run on available computing resources. For homogenous systems, this might include the reactive metal center, large bulky ligands, and even the solvent itself. For supported metals, this might include the metal, the support, defect sites, promoters, poisons, co-adsorbates, and a solvent. This would be impossible with today's current codes, which can only handle on the order of 100-200 atoms. Instead, we typically model these systems with periodic slab or cluster approaches. As we move toward examining more complicated heterogeneous systems, other features, such as chirality, electric fields, electrostatic potentials, multifunctional materials and molecules, also become important factors in the models developed.

At a fundamental level, catalysis involves active sites responsible for breaking and making chemical bonds. The impact for catalysis is at a much higher level, however, whereby it dictates the behavior of a chemical plant. The ability to bridge each of these areas requires a seamless integration of methods to describe the fundamental events and reactivity over all of these length and time scales. These scales may be categorized as (1) electronic, (2) molecular, (3) mesoscale, (4) macroscale and (5) systems analyses. Calculations must move smoothly from the realm of quantum chemical methods to molecular-level simulation of adsorption, reaction, and diffusion processes on and with zeolites, metal oxides, metals and metal sulfides.

Success in this regard will require accurate and realistic interatomic potentials and force fields, and more robust molecular simulation tools. A good fraction of the catalytic processes occur over transition-metal-containing materials. In moving beyond quantum mechanics, it becomes more difficult to describe these systems because of the lack of accurate and flexible force fields. Transition metals open up a whole range of catalytic chemistry because of the nature of d-electrons and orbitals. However, this change also makes these systems much more difficult to explain with classical force fields or potential models. This becomes especially difficult for modeling systems that undergo redox or have mixed oxidation states. Many of the industrial labs pointed to the need for robust force-fields to model transition-metal atom systems. Models should be able to handle reactivity, changes in oxidation states, charge transfer, and induced polarization.

More robust molecular simulation tools will also be required to simulate multicomponent adsorption and diffusion. In addition, much of the work on molecular simulation that has been performed to date has separated reactive and diffusive time scales. However, there is a wide range of systems where modeling both of these processes simultaneously is important. The disparate time scales that exist between diffusion and reaction processes limit our ability to model both processes simultaneously. Although there has been much discussion and even some limited application of coupling both classical molecular simulation with quantum chemical methods, there currently exists no universal code that would enable industrial researchers the ability to easily carry this out.

Application Needs in Catalysis

The needs and limitations that are more specific to homogeneous and heterogeneous catalysis are presented in Table 8.5. They include de novo design and high-throughput screening capabilities, improved molecular simulation tools, and the ability to treat more complicated bio/chemical/materials/engineering systems.

Table 8.5
Future Directions for Molecular Modeling Efforts in Catalysis

De novo design and high-throughput screening	Methods, library development, informatics, controlled syntheses
Improved molecular simulation tools	Multicomponent adsorption and diffusion, disparate time and length scale for reaction and diffusion, convergence of classical and QM methods
More complicated bio/chemical/engineering systems	Enantioselective catalysis, enzyme catalysis, fuel cells, energy storage and generation

Modeling of catalysis has typically mirrored drug-modeling developments with a lag time of about 5 to 10 years. Much of what is happening in the area of pharmaceuticals modeling is the application of molecular modeling to *de novo* design and combinatorial efforts. Molecular modeling is currently being used to develop structure-activity relationships, to develop libraries that relate the properties of specific probe species to their potential features, and to generate relevant information for informatics development and controlled synthesis strategies, which are longer-term goals.

Similar efforts have been initiated in industry in the search for new homogenous catalytic materials. Modeling heterogeneous systems is much more difficult. Library development and informatics needs will require different strategies and therefore the role of modeling will be different. It is highly likely, however, that the development and application of structure-reactivity relationships will be important. This is an area where modeling can certainly have a great impact.

Catalytic chemistry is controlled by the reaction mechanism. The ability to design active and selective catalysts requires understanding adsorption, diffusion, and the myriad surface reactions that ultimately make up the catalytic mechanism. Reliable energetic and kinetic predictions are critical for modeling efforts. Although accuracy is desired, even the ability to predict simply the correct trends is important. First-principles methods, as discussed above, have clearly reached the stage where they are beginning to be useful for computing relative trends and reliable energetic information of well-defined systems.

The ability to design more active and selective homogeneous and heterogeneous catalysts requires establishing the nature of the transient intermediates and their reactivity. Much of the current effort, at least for homogeneous modeling, is focused on elucidating catalytic mechanisms and kinetics. Unfortunately, such work has necessarily ignored the effects of changes in the catalytic structure due to processing. Catalyst life is a critical concern for operation and can indeed be a breaking point for whether or not a particular technology will be pursued. There are currently no fundamental mechanisms that could be used to describe and model catalyst deactivation and catalyst poisoning. This is an area that is vitally important, especially for heterogeneous reaction systems.

As we move into the 21st century, many of the expected breakthrough technologies will require advanced and complicated coupling of biological, electronic, chemical, and engineering components into materials development and device manufacture. This includes fuel cells and advanced energy storage devices, biocatalytic materials, enantioselective and enzyme catalysts. Modeling in these areas will require first-principles algorithms that can handle much larger systems, as well as the ability to handle chemical, biological, and engineering complexes together. It will be necessary to build bridges, in a seamless fashion, between electronic, molecular, microscale, mesoscale and macroscale systems approaches.

Implementation Needs in Catalysis

A common set of more specific needs also arose from the discussions with industry. Most of these are covered in Chapter 9. Comments are outlined here that were made specifically by the groups trying to model homogeneous and heterogeneous catalysis.

Accessibility to the source code was the first issue. Source code is restricted for much of the software that industry uses. This prevents users from improving algorithms and advancing their applications.

Second, calculations are typically carried out on various computational platforms. Researchers often have considerable difficulties in porting codes to different architectures. Improved portability of codes would enable users to take full advantage of available resources.

User-friendly interfaces would facilitate use of the codes by a broader group of scientists and engineers, and also facilitate the more experienced users' jobs.

A common source of frustration that plagues many companies is inconsistency between subsequent versions of codes. As new versions of code are released, the potentials and algorithms used may be changed or discontinued altogether. This can require that industry redo many of the calculations that had been performed with a previous version just to keep the results consistent or, in the worst-case scenario, cause them to abandon calculations entirely.

Finally, most companies the WTEC panel visited indicated a desire to calculate a greater set of properties.

Elements for Success

Key elements behind the success of modeling catalysis in industry include: (1) close integration of experiment, theory, and modeling; (2) team efforts; (3) education; and (4) partnering.

The industrial computational chemists and molecular modelers were typically found in either catalysis or modeling groups. There was roughly an even mix among the industrial sites visited. Both groups cited the close integration of theory and experiment as key to their successes. Some of what appeared to be the more successful programs were those in which modelers were directly coupled with experimental groups in catalysis. This enabled them to interact more frequently and become more closely aligned with the overall program in the course of daily interchanges of ideas.

The more successful catalysis efforts involve a team approach that closely couples catalyst synthesis, catalyst characterization, catalytic kinetics, reaction engineering, inorganic chemistry, and molecular modeling to elucidate an understanding of the mechanism and to develop new catalyst leads. This was a view shared by nearly all the industrial participants. Molecular modeling is thought to provide valuable information when it is tied directly to the experimental program.

Educating both management and other non-experts on what modeling can and cannot do has been another key element to the success of modeling. Education has also been important for promoting the use of computational chemistry and molecular modeling by non-experts in industry.

As discussed earlier, industry has begun to take full advantage of partnering with either academia or national laboratories in order to expand developments or strengthen applications. These efforts are usually fairly modest. Many of the U.S. companies would like to see the government take a more active role in promoting academic-industrial efforts. This includes improved mechanisms for funding academic, industrial, and national-laboratory collaborations. The NSF GOALI program was seen as a step in the right direction, limited mostly by the difficulty of moving people.

FUTURE DIRECTIONS

Looking into the future, it is apparent that computational power and faster algorithms will enable analysis of much larger systems and solution of much more difficult problems. The area of drug design helps to benchmark what will likely be feasible for catalysis efforts.

The greatest impacts will come from advances in both QM and molecular simulation that begin to couple the two for solving more realistic environments. How are catalytic activity, selectivity and lifetime all affected by changes in the atomic structure of working materials? This will require a detailed understanding of the electronic and atomic structure and how they change with changes in processing. For supported metal catalysts, the reaction environment would include the metal, support, defect sites, promoters/poisons, multimetallics, intermolecular interactions, and solvents. A seamless integration of methods from first-principles to complex reactive computational fluid dynamics models will ensure that the chemistry, kinetics, and fluid flow profiles are consistently integrated in terms of establishing large realistic reaction engineering and process models.

One of the main areas cited by industry as crucial for success was the close tie with experimental efforts. Looking to the future, methods should become even more tightly integrated. One could easily envision molecular simulations or quantum chemical calculations being directly accessed in the analysis of an experimental characterization. Such coupling can already be seen in some form by looking at how mass spectrometers work. The mass spectrometer ionizes and fragments the molecules that enter the chamber. Data are compared against a series of different libraries and regressed against the patterns to weight the goodness of fit. The resulting output offers patterns of the actual experiment as well as those from likely candidates for the users comparison.

It is possible to envision that subsequent to an experimental run, a quantum chemical calculation or molecular simulation would be performed to predict a possible outcome. The experimental and simulation results would be compared then in order to establish the likelihood that the two are the same. This has already begun to occur in the analysis of adsorption pore size distributions. Some of the codes that are currently sold with adsorption analysis equipment carry out molecular density functional simulations in order to establish the pore size distributions.

One could envision the coupling of FTIR, HREELS, or Raman with quantum chemical codes in order to perform similar on-line analysis of the infrared or Raman spectra. Another possibility is coupling EXAFS and structural optimization. Other areas might include *ab-initio*-based kinetic Monte Carlo simulation of temperature-programmed desorption or of Scanning Tunneling Microscopy. The tools currently exist to carry out such combinations on an individual material basis. Generalizing these efforts would be an important step toward integrating theory, simulation and experiment.

As models of the reaction environment become more complete, it becomes possible to move from the traditional mold—using computation for analysis—to developing design approaches. It should become possible to identify a host of possible metal centers, ligands, and solvents that lead to optimal catalytic properties for homogenous systems. Likewise, it should become possible to identify optimal designs of composition and morphology and size distribution of supported multimetallic particles, as well as metal-support features that govern heterogeneous catalytic systems.

If the dream of computationally aided design is ever to become real, modeling catalyst synthesis will be necessary as well. Expert-systems approaches to chemical synthesis could markedly influence the way in which new leads are identified. Such an approach would couple well with the major shifts toward high-throughput screening in the rational synthesis of new materials.

In summary, industry has currently begun the change from analysis to design for homogeneous systems, and the same change will occur for heterogeneous systems—both aided by molecularly based modeling.

REFERENCES

Aagaard, O.J., R.J. Meier, and F. Buda. 1998. Ruthenium-catalyzed olefin metathesis: a quantum molecular dynamics study. *J. Am. Chem. Soc.* 120, 7174-7182.

Barrett, P.A., R.H. Jones, J.M. Thomas, G. Sankar, I.J. Shannon, and C.R.A. Catlow. 1996. *Chemical Communications* 2001-2002.

Bell, A.T., M. Boudart, B.D. Ensley, D. Estell, R.H. Grubbs, L.L. Hegedus, L.E. Manzer, J.A. Rabo, J. Rebek, J. Roth, G.A. Somorjai, and V.W. Weekman. 1992. *Catalysis Looks to the Future.* National Academy Press.

Bell, A.T., E.J. Maginn, and D.N. Theodorou. 1997. Molecular simulation of adsorption and diffusion in zeolites. In *Handbook of Heterogeneous Catalysis.* VCH Publ.: Germany; Vol. 3; pp 1165-1188.

Bell, R.G., D.W. Lewis, P. Voigt, C.M. Freeman, J.M. Thomas, and C.R.A. Catlow. 1994. Computer modeling of sorbates and templates in microporous materials. In *Zeolites and Related Microporous Materials: State of the Art 1994*; Vol. 84; pp 2075-2082.

Besenbacher, F., I. Chorkendorff, B.S. Clausen, B. Hammer, A. Molenbroek, J.K. Norskov, and I. Stensgaard. 1998. *Science 279,* 1913.

Blaszkowski, S., and R.A. van Santen. 1996. *J. Am. Chem. Soc.* 118, 5152-5153.

Catlow, C.R.A. 1997. Computer simulations. In *Handbook of Heterogeneous Catalysis.* VCH Publ.: Germany; Vol. 3; pp 1149-1165.

Cheong, M., R. Schmid, and T. Ziegler. 2000. Density functional study of the migratory insertion step in the carbonylation of methanol catalyzed by [M(CO)2I2]- (M=Rh,Ir).

Dahl, S., A. Logadottir, R.C. Egeberg, J.H. Larsen, I. Chorkendorff, E. Tornqvist, and J.K. Norskov. 1999. *Phys. Rev. Lett. 83,* 1814.

Eniricerche S.p.A., 1994. Italian Patent Appl. MI94/A 002037.

Farrauto, R., and C.H. Bartholomew. 1997. *Fundamentals of Industrial Catalytic Processes.* Blackie Academic & Professional: London.

Frash, M., and R.A. van Santen. 1999. *Topics in Catalysis 9,* 191-205.

Frenkel, D. and B. Smit. 1996. *Understanding Molecular Simulation: From Algorithms to Applications.* Academic Press: San Diego.

Froese, R.D.J., and K. Morokuma. 1999. *Chemical Physics Letters 305,* 419-424.

Fusco, R., L. Longo, L. Gila, and F. Cambisi. 1996. Catalyst with improved activity for ethylene (co)-polymerization. In *EP0741146A1.*

Gonzales, N., A.T. Bell, and A.K. Chakraborty. 1997. *J. Phys. Chem. B 101,* 10058-10064.

Govind, N., Y.A. Wang, and E.A. Carter. 1999. *Journal of Chemical Physics 110,* 7677-7688.

Hackmann, M., and B. Rieger. 1997. *Catech 2,* 79-92.

Hagen, J. 1999. *Industrial Catalysis: A Practical Approach* ; Wiley-VCH: Weinheim.

Hammer, B., and J.K. Norskov. 1997. Theory of adsorption and surface reactions. In *Chemisorption and Reactivity on Supported Clusters and Thin Films,* Pacchioni, R.M.L.a.G., Ed. Kluwer Academic Publ.: The Netherlands, pp 285-351.

Hammer, B., and J.K. Norskov. 2000. *Adv. in Catal. 45,* 71-129.

Hansen, E.W. 2000. Monte Carlo simulation of heterogeneous catalytic reaction systems. Ph.D. Dissertation, University of Virginia.

Hansen, E.W., and M. Neurock. 1999. *Chem. Eng. Sci. 54,* 3411-3421.

Hansen, E.W., and M. Neurock. 2000. *J. Catal.* (in Press).

Haynes, V., J. Armor, A. Bell, J. Ebner, B. Goodall, R. Grubbs, J. Lerou, A. Manheim, T. Marks, C. Murchison, B. Smart, T. Vanderspurt, and B. Warren. 2000. *Vision 2020 Catalysis Report.* Department of Energy, OIT.

Hermann, K., M. Witko, and A. Michalak. 1999. *Catal. Today 50,* 567-577.

Kang, H.C., and W.H. Weinberg. 1995. *Chem. Rev. 95,* 667-676.

LaPointe, R.E., J.C. Stevens, P.N. Nickias, and M.H. McAdon. U.S. Patent Number 5721185: *Homogeneous olefin polymerization catalyst by abstraction with Lewis acids.*

Lohrenz, J. 1998. Quantenchemie in der chemischen Industrie: Von der Berechnung von Molekülstrukturen bis zur Simulation von chemischen Reaktionen. Höchstleistungsrechnen in der Chemie, Forschungszentrum Jülich,; http://www.zam.kfa-juelich.de/compchem/proceedings/Lohrenz/lohrenz.html.

Margl, P., L. Deng, and T. Ziegler. 1999. *Topics in Catalysis 7,* 187-2000.

McAdon, M.H., J.T. Patton, P.N. Nickias, R.B. Shankar, F.J. Timmers, B.W. Kolthammer, D.D. VanderLende, and S.M. Ueligger. U.S. Patent Number 5965756: *Fused ring substituted indenyl metal complexes and polymerization process.*

Meier, R.J. 1999. Applications of ab-initio molecular dynamics simulations in chemistry and polymer science. In *Computational Molecular Dynamics: Challenges, Methods, Ideas, Lecture Notes in Computational Science and Engineering,* M. Griebel, D.E. Keyes, R.M. Nieminen, D. Roose and T. Schlick, eds. Springer-Verlag.

Meier, R.J., G.H.J. van Doremaele, S. Iarlori and F. Buda. 1994. Ab-initio molecular dynamics study of metallocene catalyzed ethylene polymerization. *J. Am. Chem. Soc.* 116, 7274-7281.

Neurock, M. 1997. *Stud. in Surf. Sci. and Catal. 109,* 3-34.

Neurock, M., and. E.W. Hansen. 1998. *Comp. Chem. Engng. 22,* 1045-1060.

Neurock, M., and A.T. Bell. 1997. Future directions for theoretical catalysis, homogeneous, heterogeneous, and surface reactivity. National Science Foundation and Department of Energy.

Neurock, M., and R.A. van Santen. 1994. *J. Amer. Chem. Soc. 116,* 4427-4439.

Nicholas, J.B. 1999. *Topics in Catalysis 9,* 181-189.

Nickias, P.N., M.H. McAdon, J.T. Patton, B.P. Friedrichsen, J. Soto, J.C. Stevens, and D.D. VanderLende. U.S. Patent Number 6015868, *Substituted indenyl containing metal complexes and olefin polymerization process.*

Pacchioni, G. 1995. *Heter. Chem. Rev. 2,* 213-228.

Payne, M.C., M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos. 1992. *Reviews of Modern Physics 64,* 1045-1097.

Sandre, E., M.C. Payne, and J.D. Gale. 1998. *Chemical Communications* 2445-2446.

Sankar, G., J.K. Wyles, R.H. Jones, J.M. Thomas, C.R.A. Catlow, D.W. Lewis, W. Clegg, S.J. Coles, and S.J. Teat. 1998. *Chemical Communications* 117-118.

Sankar, G., J.M. Thomas, C.R.A. Catlow. 2000. *Topics in Catalysis* 10, 255-264.

Sauer, J. 1989. *Chemical Reviews 89,* 199-255.

Sauer, J. 1994. Structure and reactivity of zeolite catalysts—atomistic modeling using ab-initio techniques. In *Zeolites and Related Microporous Materials: State of the Art 1994*; Vol. 84; pp 2039-2057.

Sauer, J., P. Ugliengo, E. Garrone, and V.R. Saunders. 1994. *Chemical Reviews 94,* 2095-2160.

Schreck, P.d.w.D. J. 2000. U.S. Patent.

Sierka, M., and J. Sauer. 2000. *Journal of Chemical Physics 112,* 6983-6996.

Smith, J.A. 1999. Catalyst composition for the polymerization of olefins. In *EP0889061.*

Stampfl, C., M.V. Ganduglia-Pirovano, K. Reuter, and M. Scheffler. 2001. *Surf. Sci. 500.*

Stich, I., J.D. Gale, K. Terakura, and M.C. Payne. 1999. *Journal of the American Chemical Society 121,* 3292-3302 .

Teunissen, E.H., A.P.J. Jansen, R.A. van Santen, R. Orlando, and R. Dovesi. 1994. *J. Chem. Phys. 101,* 5865-5875.

van Santen, R.A., and. M. Neurock. 1998. *Theory of Surface Chemical Reactivity*, Ed.; VCH-Wiley: Berlin.

van Santen, R.A. 1997. *Journal of Molecular Catalysis a-Chemical 115,* 405-419.

van Santen, R.A., and G.J. Kramer. 1995. *Chemical Reviews 95,* 637-660.

van Santen, R.A., and M. Neurock. 2000. *Molecular Modeling of Heterogeneous Catalytic Reactivity*, Encyclopedia of Catalysis.

van Santen, R.A., and M. Neurock. 1995. *Catal. Rev.—Sci. Eng. 37,* 557-698.

Vanderbilt, D. 1990. *Phys. Rev. B 41,* 7892.

Venkataraman, P.S., M. Neurock, and G.W. Coulston. 1999. *J. Phys. Chem. 103,* 8973-8983.

Vreven, T., and K. Morokuma. 2000. *Journal of Chemical Physics 113,* 2969-2975.

Wang, X.-G., A. Chaka, and M. Scheffler. 2000. *Phys. Rev. Lett. 8,* 3650-3653.

Whitten, J.L., and H. Yang. 1996. *Surf. Sci. Rep. 24,* 59-112.

Witko, M., K. Hermann, and R. Tokarz. 1999. *Catal. Today 50,* 553-565.

Woo, T., P.M. Margl, L. Deng, L. Cavallo, and T. Ziegler. 1999. *Catal. Today 50,* 479-500.

CHAPTER 9

INFRASTRUCTURE ISSUES FOR APPLYING MOLECULARLY BASED MODELING[2]

Ellen B. Stechel

INTRODUCTION

This chapter presents a management point of view. The intent is to represent what the panel learned about attributes in a research environment that support or obstruct the successful application of chemical and materials modeling.

There should be little question that the opportunities are rich. There is growing recognition and acceptance of the value of molecular and materials simulation in National Laboratories, in academia, and in industry. Currently in chemical and materials industries approximately 1-5% of the research staff is computational; however, some expect that number to grow to 25% in the next ten to fifteen years—even to 50%. Expanded funding and programmatic support are fueling some enthusiasm nationally; e.g., the ASCI program (Accelerated Strategic Computing Initiative of the U.S. Department of Defense, http://www.lanl.gov/projects/asci/asci.html) and the IT^2 or ITR Presidential initiative (Information Technology Research, http://www.itr.nsf.gov/). However, the impact of these initiatives on molecular or materials applications remains uncertain.

In addition, computational resources are becoming increasingly affordable, although the cost of expertise and software remain considerable. Computational methods and algorithms are evolving rapidly; the field remains intellectually dynamic. There is some attempt to make user-friendly, widely available codes, and the assumption is that the enabling tools will continue to improve over time.

As a consequence, the prevailing view is that the most aggressive companies should eventually develop a definitive competitive advantage. Some would say that companies already are deriving a notable competitive advantage from their molecular-scale modeling, for example through developing broad, defensible patent positions.

Even given the impressive and immense technical advances discussed in other chapters, successful application in an industrial or mission-oriented setting requires, in addition, a large number of other non-technical attributes—attributes not always considered. Some of these attributes have to do with people, including the following:

- Skill-sets
- Critical mass

---

[2] Copyright reserved by Ford Motor Company. Reproduced by permission.

- Quality of the colleagues
- Education and training of experts and non-experts
- Respect amongst the disciplines
- Alignment of rewards and recognition

A second set of attributes pertains to setting and achieving relevant goals:

- Maintaining realistic expectations
- Appropriate measures of success
- Decisive and informed value-added management with vision for the future
- Articulate and vocal management champion (preferably more than one)
- Compelling and reasoned research strategy
- Research programs with transparent and understandable connection to the mission and goals of the organization

A third set of attributes concerns innovation and cross-fertilization, e.g.:

- Integrated and relevant research programs that molecular and materials modelers can be an essential part of
- Availability of experimentalists to effectively interact with
- A environment that supports a collaborative and teaming multi-disciplinary culture
- External collaborations, including some proximity to centers of excellence

Lastly, success depends on having strong internal support systems, including both hardware (support and adequate resources) and software (availability of user-friendly, robust, flexible, adaptable, affordable, and reliable code).

To organize our findings, we first discuss some background information about the nature of problem solving in industry, which is not necessarily equivalent to problem solving in an academic setting. It is important to understand the differences. We next discuss the meaning of successful application of chemical and materials modeling in an industrial setting. This is a topic of ongoing debate, so primarily we will summarize the elements of discussion. There is, nonetheless, a great need to clarify this issue. Lastly, we will discuss those attributes (beyond technical) that are necessary to achieve success.

There is every reason for optimism and to believe that recent advances in molecular and materials modeling place us on the verge of something revolutionary. Nevertheless, it is necessary to temper that enthusiasm because it is equally unmistakable that there is some distance yet to travel.

PROBLEM SOLVING

It is important to realize that industry researchers are primarily "users" of this technology. The WTEC panel saw some exceptions, notably in Japan (see the Doi Project and JRCAT site reports, for example). In those cases, the general motivation seemed to be to reduce dependency on the U.S.-dominated commercial software market. Nevertheless, non-software industries would prefer to be able to buy software tools, and not have to invest enormous resources in them. The primary reason is simply the time scale for getting results; time is critical to competitive advantage. Hence, it is important to understand the needs of the users. Those needs include computational robustness, reliability, efficiency, ease of use for creating the input and for interpreting the output, and ease of visualizing and communicating results.

As indicated above, industrial and mission-oriented problem solving requires a different approach than typical academic research. To deal with the complexity of industrial or societal problems, one must deal with realistic, not overly idealized, systems. One must be sensitive to the time available to solve real problems, as timeliness is heavily valued. Industry requires a multidisciplinary approach and working in teams. Industry

generally requires access to and understanding of the full range of methods available to solve a given problem. Academics, on the other hand, are able to select problems to match available tools and expertise without limiting their opportunities for success. Academic research can focus on developing new methods through proof of principle. This, however, is not always commensurate with a needed focus on testing and validation. Academic research follows the line of what makes for the best thesis, i.e., "educate the student." While industrial researchers also tend to tailor their problem choices to the available tools, this regrettably *is* a limitation to success. To achieve success, the industrial researcher must be able to transfer knowledge—in a timely fashion—that will lead to the solution of real problems.

Because industrial researchers are primarily users of simulation technology, it is necessary that there be a number of vehicles for transferring this technology into industry. Few industries use all these mechanisms. Some of them are used by only a small minority of companies. Various examples include licensing software, creating small startups, purchase of commercial software, and use of software consortia (MSI, GdR). All these vehicles, and probably others, serve to move simulation tools into widespread use. Another class of vehicles moves the technology into industry but not necessarily into widespread use. The basis for this rests on the idea that to transfer technology most effectively one transfers the people possessing the knowledge. Examples include industry/academic collaborations, contract research, industrial hiring of personnel who bring research code and experience, exchange of personnel (visiting scientists in industry), industrially supported postdoctoral positions (either in-house or through university collaborations), summer intern programs, and sabbaticals in academia. One model worth mentioning is that of the Materials and Process Simulation Center (MSC) at Caltech (http://www.wag.caltech.edu/), where one of the primary goals is to "facilitate the technology transfer of advances in atomistic simulations of materials and molecules from university research laboratories to industrial practice." This model has evolved through extensive consulting and research activities between Professor Goddard and his team and various companies over the years (including Shell Development Corp., Allied Signal, ICI, General Electric, Molecular Simulations Inc., General Motors, SOHIO, Nippon Steel, Dow Chemical, and Xerox Corp.).

THE MEANING OF SUCCESSFUL APPLICATION

If implemented correctly, modeling involves comparatively little human effort relative to the amount of information generated—compared to other means. Nonetheless, the measure of success is impact. There are often many ways to generate impact, some tangible, and some intangible. An example of intangible impact is guiding and interacting with experimentalists to enhance effectiveness in the design of experiment and in deriving inferences from experimental results. Judicious use of simulation can provide a rationale for the elimination of experiments that would not provide useful information or a fruitful avenue for inquiry.

In order to organize our conclusions on how researchers can derive significant value from computation, we classify simulation into five basic modes: discovery, calculation, benchmarking, validation, and prediction. Each provides an avenue to impact.

Discovery. The discovery mode seeks to find new relations and paradigms for complex behavior. In addition to the possibility of generating intellectual property, this can also provide insights to guide experiment (including what not to do), which in turn can lead to substantial, but intangible payoffs. In addition, modeling can provide the basis for simpler, faster QSAR or QSPR models (Quantitative Structure-Activity Relations or Quantitative Structure-Property Relations); simple usable models and relationships are tangible deliverables. Another plausible outcome is broadening the claims in a patent application that might have originated from an empirical observation.

Calculation. In the calculation mode, one computes or determines parameter input for simplified mathematical models. This offers a tangible deliverable by providing needed model input that might be difficult or impossible to measure experimentally.

Benchmarking. Benchmarking is a worthwhile use of simulation in disciplined testing of approximations or assumptions in simplifying mathematical models. When done appropriately, this can build confidence and credibility for the models. It also can supply a characterization of reliability and accuracy. Indeed, we will

reach a point when modeling and experimental labels on data will become meaningless. That time will come when computational and experimental data become equally reliable.

Validation. The definitive way to build confidence and develop measures of reliability is through careful comparison between simulation and experiment. This creates the necessity for effective use of the validation mode. In addition to determining credibility and reliability, one can also gain greater knowledge or information through interpreting experimental analytical data. This synergy between experiment and simulation has the effect of delivering far greater value and more information than the sum from each working independently.

Prediction. With simulation, one can predict a wide range of difficult to observe behavior of complex molecular and materials processes and phenomena. This can generate intellectual property, through prediction of either function or properties in a specific or general class of molecules or materials. Again, the greatest value typically arises from the synergy of working with complementary experimentation to verify hypotheses inferred from computational results or to generate hypotheses to verify with computation.

The precise metrics for success depend on whether the goal is for simulation to remain a specialist tool or for it to become a ubiquitous tool. Some believe that success will occur when the bench chemists and materials scientists are equally adept in using computational tools as experimental tools. This is the case to some extent with CFD (computational fluid dynamics) and CAE (computation aided engineering.) Others are, nonetheless, comfortable with molecular and materials modeling remaining as specialist tools. There are precedents. It still requires significant expertise to gain the greatest effectiveness from materials and molecular experimental characterization. When used as a specialist tool, potential impacts of modeling include development of insight to guide and interact with experiment, delivering simple models or design rules, interpreting observational data, cost savings (e.g., by eliminating expensive experiments), and intellectual property development through generating defensive publications and patents (including broadening claims and defending patent positions). The service mode of operation (analogous to analytical labs) has some utility, and some industries operate in this mode. Then the measurement for success is paying customers.

NECESSARY ATTRIBUTES FOR SUCCESS

Development of the Necessary Human Resources

Success typically depends on (1) the right people (2) working on the right problem (3) with the right tools. This section reports our findings on the part of the equation that focuses on the right people. Even to recruit and retain the right people, the organization must have a high quality of staff to provide a stimulating and inspiring environment. It must align its rewards and recognition to provide appropriate motivators and feedback to reinforce desired behavior and successes. To have the right molecular and materials modelers, there needs to be critical mass and strong interaction amongst the modelers, without isolating those modelers. The people either must come in with the necessary skill sets or must be able to quickly develop those skill sets. We focus here on the development of the important skills for researchers in industry.

In an ideal world, development of these skills would start at the K-12 level, where it is important to capture and hold onto the imagination of all students. It is worth noting that visualization using simulation tools can be a great aid in this endeavor. This is the time to develop the drive and the motivation for life-long learning. The undergraduate level is the time to learn how to learn, how to effectively use questions to advance understanding, and equally importantly to build the foundations; e.g., in math, computer science, chemistry, physics, and biology. Graduate school is the time to build the skill set to succeed. Herein lie some potential issues.

Success means different things in academia, national and government labs, and industry. In academia, success means educating to the standards of excellence as defined by other scientists with similar interests and similar. By contrast, in industry impact on the bottom-line carries the day; this requires more than technical talent. It requires communication skills, in particular the knack to excite and relate to a wide range of people with very different interests and skill sets. Another important skill or talent is innovation, the

ability to dream up and implement a non-obvious idea. One needs the ability to recognize what knowledge might be useful to someone else and be able to deliver it while that person retains an interest. It is essential to appreciate that results that are more accurate can often be less valuable than results that are less accurate but arrived at more quickly. This is especially true when the added accuracy does not affect how the person will use the knowledge. Like the academic researcher, the industrial researcher must have the ability to generate intellectual content (depth) and must have a thirst for life-long learning (breadth). Lastly, the researcher must possess a drive to succeed and to make a difference.

Development of the Necessary Support Infrastructure

There are four elements enabling scientific computing: (1) theoretical and mathematical formulation, (2) algorithmic design, (3) software development and implementation, and finally (4) compatibility with system architecture. Typically, these are not well integrated. The consequence is that development of these enablers are rate limiting. Many scientific and engineering advances still must wait. Nevertheless, the advances in both hardware and software are enabling significant applications, as already highlighted in the previous chapters.

With respect to hardware, the WTEC panelists found that many industries have invested in significant hardware resources to support scientific and engineering computation. Typically, we found SGI Origins—up to 64 processors. We also found that most organizations had high-end desktop workstations for their modelers. However, hardware is changing rapidly. Supercomputing is moving in the direction of commodity-based processors, and therefore it will soon become widely available and affordable. It had been available only at national labs, defense labs, and NSF supercomputing centers (which are now lagging behind). We now find most organizations, at least, talking about building a Linux-based cluster of either Intel Pentium or Compaq-Alpha processors in the near future. At Motorola, we found already in existence a group-managed, Intel-based, 256-processor cluster. There exists an impression that hardware is no longer a limitation; however, affordable, available resources are far from "infinite" and far from inexpensive. The potential issue here is that managers in most organizations might not really have this in their plan for investing in capital resources or they might not fully understand the changing landscape.

Counting processors and potential computing power is not the correct evaluation metric. The determining factor for the value of the system is effectiveness, not number of processors or cost per processor or peak flop rate (http://www.nersc.gov/aboutnersc/esp.html). What is most important is the amount of useful calculations the system can support. The ultimate measure would be the "volume" of science accomplished—but no one knows how to measure that. Since what you measure is what you get, there remains a critical need for effective performance indicators. For example, consider increasing efficiency from 80% to 90%. 256 processors running at 90% is equivalent to 288 running at 80%. In other words, one gets 32 extra processors for no added power, space, or monetary cost. Hence, effective system determination is extremely important in our cost-determining world, both before and after purchase. The peak CPU measure was useful when the CPU was the only major expense. It persists only because it is simple, although quite uninformative. In general, it is hard to measure the truly important things, so simple measures endure, whether they are appropriate or not, as long as no one can define anything better. NERSC (http://www.nersc.gov, DOE civilian supercomputing center at Lawrence Berkeley National Laboratory) is making a serious attempt to create such a metric. However, it is not evident that anyone in industry is thinking in terms of system performance metrics.

Another issue we consistently found in industry the lack of a clear understanding between researchers and their central computing support. There is insufficient appreciation that managing computer resources for intensive computation has unique needs relative to that for desktop administrative and business applications. This is in stark contrast to the management of computer resources at national laboratories and in academic institutions. However, even in national laboratories and in academic institutions, there is a continuing dichotomy between central resources and distributed resources. Nonetheless, in general there seems to be a clear, real distinction between computers for computation and computers for administrative purposes. In industry, we observed many research groups with a preference to manage their own computer resources, despite what it costs them in their own human resources.

Software is the one area that generated a great deal of discussion, even controversy, during the WTEC site visits. Clearly there are many unmet needs, but very few answers as to how to meet those needs. Until we arrive at a point where generating software is both cheap and easy, these very hard issues will persist. For industries, the discussion typically goes along the lines of make or buy. Generally, in industry, there is a strong preference to buy it, if possible, and to make it only if it is unavailable for purchase and meets a critical and recognized need. One can observe some regional differences. Specifically, in the United States very rarely does one see an industrial researcher developing either chemical or materials modeling software. In Europe there is some industrially developed software (Dissipative Particle Dynamics by Unilever, for example), but this approach is not extensive. In Japan, in-house software development is quite extensive; the primary reason cited is the desire to be independent from U.S. dominance of the commercial software market.

There was a great deal of discussion about "academic" code as opposed to "commercial" code. Commercial code typically starts as academic code. Nonetheless, considerable effort is necessary to make an academic code commercially viable and usable. In molecular and materials modeling, widespread commercial codes include those developed by MSI, Gaussian, Schroedinger, CPMD, and VASP; each operate with a different business model. There also exist some public cooperative efforts—ESF-PsiK (http://psi-k.dl.ac.uk/), NIST-CTCMS (http://www.ctcms.nist.gov/), DOE-CSMN (http://cmpweb.ameslab.gov/ccms/)—with the goal of developing and disseminating community code. These efforts foster collaborations and develop state-of-the-art freeware tools and techniques.

The MSI Consortia represent one mechanism and business model for developing commercial code for industrial customers. The model recognizes that close partnership with industry is essential to develop software and services that help to solve critical research problems. MSI uses a partnering mechanism, with the objectives of channeling industrial knowledge into the design of new capabilities and making simulation expertise available to commercial researchers. Consortia are groups of industrial researchers, academic experts, and MSI scientists, focused on developing, validating, and applying simulation to a particular research area. Membership provides software, formal input (by democratic vote) to MSI product plans, regular meetings to share experiences and ideas, and dedicated application support. There are currently four materials science consortia covering polymers, catalysis, formulations (knowledge management, decision support, and informatics,) and crystallization (for pharmaceutical development).

Since its inception over ten years ago, MSI views the consortium model as being highly successful, particularly in establishing simulation in new application areas. MSI claims that consortia provide nuclei from which wider use grows. From the MSI vantage point, members enjoy the competitive advantages of early adapters—a technology lead, a greater knowledge base, and better long-term return on investment. The panel also heard some differing viewpoints from industry. Members get the software, but some member companies were uncertain what else they really obtained. One consortium member stated his perception that "we used to be collaborators, but now we're customers."

The WTEC study indicates that there has yet to be demonstration of success for the consortium model—at least something is missing. This indicates at least a need for credible metrics to evaluate the effectiveness of the model. Some such metrics would include measures of code reliability, of timeliness in incorporating state-of-the-art methodologies, and of portability.

Linear scaling provides an example that highlights some of the challenges. Repeatedly, industrial researchers cited the promise of linear scaling as one key enabler. Linear scaling in quantum calculations is an algorithm with CPU and memory requirements that scale linearly with system size as opposed to cubic scaling (as in DFT). This would make it conceivable for some calculations to be in the 10,000-50,000 atom range. Is this a solved problem? The answer is clearly not, if one uses the metric that it will be a solved problem when it is in widespread reliable use in several quantum-chemical software packages available to industry. Nonetheless, the progress is certainly promising. There are numerous proofs of principle, concepts, some tight-binding applications, and few if any accurate first-principles DFT. So why is it not a solved problem? It is because there are numerous remaining technical issues—"the devil is in the details." There are barriers for academics to pursue beyond this point. It is not easy to insert into existing codes without major rewrites or until there is extensive modularization, standardization, and algorithm design science. This is at a stage still beyond the scope of software companies. Therefore, this precludes general implementation into

commercial codes. We use this as an example, something often cited in industry as an expected enabler, but like many other academic advances, it remains slow to get into widespread use. Clearly, something is missing in the necessary structure or incentives to radically decrease the time or the effort that it takes to develop new software.

Additionally, considerable programming challenges exist at the hardware/software interface. The programmer still must play a substantial role to achieve performance on a system with greater than 16 processors. The development of scaleable software is another very significant challenge. It is enormously difficult to write programs that run well on 4 to 256 processors or more. Indeed, there is a sizeable and growing gap between ideal peak performance and achievable performance with scientifically useable code.

Cultivating a Culture for Innovation and Cross-Fertilization and Having a Strategy for Setting and Achieving Relevant Goals

The cultural environment and the nature of the research programs can spell the difference between success and failure. If there are no integrated and relevant research programs that molecular and materials modelers can be an essential part of, if there are no experimentalists that modelers can effectively interact with, if the environment does not support a collaborative and teaming multi-disciplinary culture, success should not be expected. Success will depend on the organization having a culture that can support an equal partnership across theory, simulation and experiment. Rarely will molecular or materials simulations single-handedly solve a problem. Success seems to depend on early involvement, which depends on having at least pockets of acceptance amongst experimental researchers. There must be communication and appreciation for what simulation can bring to the table.

It is helpful, possibly essential, to have articulate and vocal management champions. However, unrealistic expectations often lead to failure, so maintaining high but realistic expectations is critical. In addition, it is important to have proper measures of success, to have decisive and informed management with vision for the future, to have a compelling and reasoned research strategy, and to have research programs with a transparent and understandable connection to the mission and the goals of the organization. These factors determine what problems will have impact and whether there is reason to expect success.

Unfortunately, many (possibly most) experimentalists still believe that at best modeling can interpret experiment—consequently, they generally call the theorist in only when they are stuck. However, as mentioned previously, early involvement in projects is critical to success. Another critical factor for success is a strong connection to external researchers, in order to stay near the state of the art, to maintain credibility, and to derive leverage. Possibly, this is why industrial success shows a correlation with the existence of a proximate center of excellence—either a strong academic institution or a federal laboratory.

CLOSING REMARKS

A research environment is enormously complex and there are so many elements that are necessary, none of which are sufficient, to guarantee success. This chapter attempted to cover the range of attributes that affect success. Given that wide range, some points warrant re-emphasizing.

One general and significant conclusion is that exceptional results from any individual contribution are becoming and will continue to become less important to an organization. Nonetheless, all too often experimental researchers turn to computational scientists only after repeated experimental failures. Had the experimentalist consulted the computational expert, it is plausible that many of the dead ends would have been avoided. Hence, it is advantageous to include a simulation piece at the beginning of a new project.

In addition, and possibly most important, theory and simulation should never be isolated but must be closely coupled with extensive experimental capabilities in order to derive its greatest value, or in many cases any significant value. A related conclusion is that to achieve integration from fundamental science to real-world applications, the real challenge is one of coordination, integration, and communication across disciplines, approaches, organizations, and R&D sectors. Teaming, communication, partnerships, cooperation, and

leverage are key. Therefore, the skill sets of practitioners must include the ability to be interactive and communicative beyond their discipline.

A second major conclusion is that maintaining realistic, albeit high, expectations is critical to allow development time and needed cultural changes for this core competency to achieve its potential greatest successes. Considering that what an organization measures is what the organization will get, the organization should be mindful in what it asks for. Therefore, it is critical to understand that there is more than one way to positively impact organizational goals and that some of the less tangible measures are often the most important. An example of a reasonably tangible measure occurs when molecular and materials simulation has enabled either broadening or enhancing the defensibility of a patent position. Less tangible successes include guiding experiments to avoid dead ends and to gain greater knowledge from observations. In addition, since the breakthroughs will most likely arise at the interfaces and from synergy, continued undue emphasis on individual accomplishments and individual contributions would serve to stifle innovation and progress.

A third and final major observation concerns the nature of required hardware resources, which is changing rapidly and driven by relatively inexpensive, powerful commodity processors. Currently, industry is mostly at a relatively low level of exploiting parallel architectures. This is largely because the barrier to entry is still too high and commercial vendors have not optimized for various architectures. Similarly, most organizations today must choose their problems from what available software can address, as opposed to choosing problems driven purely from the application need. This is because most companies, especially in the United States, prefer to buy commercial software and only develop software internally when there is a clear proprietary or competitive advantage. However, societal impact remains limited when advances in science are slow to get into widely distributed use. There is an enormous need for all new codes to be modular and easily modified or expanded with some common protocols so "academic" advances can be rapidly incorporated into existing codes. The computer science and applied mathematics enablers exist such that the community could work to achieve a vision where the software development is no longer the rate-limiting step in advancing industrially or mission-relevant applications through molecular and materials simulation. What do not exist are appropriate incentives to achieve this "vision."

CHAPTER 10

INTERNATIONAL COMPARISON

Phillip R. Westmoreland and Peter A. Kollman

INTRODUCTION

The central purpose of this panel has been to compare and contrast the applications of molecularly based modeling around the world.

Information was gathered by interviews, from the literature, and from the panelists' professional experience. Site reports were assembled for this study from a diverse set of companies, government laboratories, and academic institutions in Europe, Japan, and the United States. Many of the locations were visited, while phone and e-mail interviews were also used. Commercial applications of molecularly based modeling include other companies and other countries and regions as well. Some of those activities are discussed in the course of this chapter.

All this information combines to form the basis for the previous chapters and for this chapter's cumulative comparisons. The panel has identified or confirmed specific regional differences and developed opinions on their causes, but we also have identified many features in common. One is that most of the technical and stylistic differences proved to be between different industries, not different regions. Second, regional differences are sometimes societally based, but more often they have originated because of the intellectual achievement of specific people or groups, only spreading gradually to other regions despite global communications and an abundance of conferences. Third, while academia is focused on the intellectual advances, successful application (or not) depends on a mix of scientific/technological and implementation issues. Finally, looking at new and expected trends, there are several surprises.

This chapter first reviews the breadth of modeling activities around the world. Next, government-sponsored activities are described and compared. Third, the activities of professional societies and other non-government, multiparty organizations are described. Finally, overall comparisons and suggested trends are offered, summarized approximately by a table of ratings.

INTERNATIONAL ACTIVITIES IN MOLECULARLY BASED MODELING

This section presents extensive tables of international activities. It is organized first by region: Europe (Table 10.1); Japan (Table 10.2); Africa, Australia, and other Asian (Table 10.3); North and South America outside the United States (Table 10.4); and the United States (Table 10.5). For a given country, commercial activities are listed first, then government-sponsored consortia and laboratories, and finally academic laboratories for countries other than the United States. Links to Web-based information are shown where possible, selecting English-language versions where there is a choice.

Table 10.1
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Europe-wide			
European Union RTN: "Quantum Chemistry of the Excited State" (7/1/1996-6/30/2000)	*Ab initio* modeling of electronic and photonic excitation, joining groups from Lund, Toulouse, Valencia, Tarragona, Leuven, Dresden, and London	Prof. Bjorn O. Roos, Lund University, Sweden	http://www.teokem.lu.se/qucex/, http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000B6AB&_app.page=show-NET.html
European Union RTN: "Potential Energy Surfaces for Molecular Spectroscopy and Dynamics" (7/1/1996-6/30/1999)	Development and application of new methods for calculating potential energy surfaces for atmospheric, combustion, and catalytic reactions; 10 groups from Italy, Denmark, Finland, France, Germany, Spain, UK	Prof. Paolo Palmieri, Bologna	http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000B71B&_app.page=show-NET.html
European Psi-k Network	Two EU Framework IV TMR networks, a Framework V RTN, CCP-9, ESF program STRUC (below)		http://psi-k.dl.ac.uk/
European Union, TMR Network: "*Ab initio* Calculations of Magnetic Properties of Surfaces, Interfaces and Multilayers" (9/1/1996-8/31/2001)	Network of groups in *ab initio* calculations on magnetic multilayers and surface magnetism; 13 universities, two research institutions (Daresbury and Jülich) and an industrial laboratory (Philips)	Dr. Walter Temmerman, CCLRC (UK)	http://psi-k.dl.ac.uk/TMR1/default.html, http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000B727&_app.page=show-NET.html
European Union, TMR Network: Psi-k project, "Electronic Structure Calculations of Materials Properties and Processes for Industry and Basic Sciences" (3/1/1998-2/28/2002)	Molecular processes on surfaces (ISMANS), f-electron systems (Univ. of Aarhus), interfaces (Queens University Belfast), magnetism (T.U. Wien), development of the LAPW method (Forschungszentrum Jülich), electronic excitations (Helsinki Univ. Technology), superconductivity (Univ. Würzburg), training programs and dissemination (Daresbury)	Dr. A. T. Svane, Univ. of Aarhus	http://www.cordis.lu/tmr/, http://www.ifa.au.dk/~svane/tmr-psik.htm, http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000B7ED&_app.page=show-NET.html
European Union RTN: "Computational Magnetoelectronics" (10/1/2000-9/30/2004)	DFT modeling of relevant interfaces, junctions, and other two-dimensional nano-structures; 10 groups including P.M. Levy of New York Univ.	Dr. Walter Temmerman, CCLRC (UK)	http://psi-k.dl.ac.uk/magnetoelectronics/default.html, http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000D2CE&_app.page=show-NET.html
European Union RTN: "Development, Validation and Application of Hybrid Atomistic Modelling Schemes" (9/1/2000-8/31/2003)	Testing different models at nanoscales for calculating electrical conductance, friction, chemically induced cracks, grain boundaries in dislocation; groups from Germany, Italy, the Netherlands, Spain, UK	Dr. Michael C. Payne, Univ. of Cambridge	http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000D3D6&_app.page=show-NET.html

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Europe-wide			
European Union RTN: "New Routes to Understanding Polymer Materials Using Experiments and Realistic Modelling" (3/1/1998-2/28/2002)	Polymer properties in terms of monomer-level structure and chain motions, studies by MD, MC, NMR, and neutron-scattering methods; groups from Belgium, France, Germany, Greece, Spain, UK	Dr. Julian Clarke, UMIST	http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000BA54&_app.page=show-NET.html
European Union RTN: "Molecular Design of Functional Liquid Crystals" (11/1/1997-10/31/2001)	Experiments and molecular simulation to predict properties of new liquid crystals; nine groups from France, Germany, Italy, Portugal, Spain, UK	Prof. Pier L. Nordio, Padua	http://www.tmr_licrys.chfi.unipd.it/, http://improving-rtn.sti.jrc.it/network/show.gx?Object.object_id=NW_VAC--000000000000BAB1&_app.page=show-NET.html
European Science Foundation, STRUC Programme: "Electronic Structure Calculations for Elucidating the Complex Atomistic Behaviour of Solids and Surfaces" (1998-2002)	Project for research and information exchange among condensed-matter electronic-structure modeling groups in Europe	Profs. V. Heine, Univ. of Cambridge, and Jürgen Hafner, T.U. Wien	http://www.esf.org/physical/pp/struc/struca.htm; http://www.esf.org/physical/pp/struc/strucb.htm
European Science Foundation, SIMU Programme: "Challenges in Molecular Simulations: Bridging the Length and Time-scale Gap" (1999-2003)	Research and information exchange among multi-scale modeling groups in Europe, workshops at CECAM; 140 European laboratories; member countries are Belgium, Denmark, Finland, France, Germany, Italy, The Netherlands, Norway, Portugal, Sweden, Switzerland, United Kingdom	Prof. Michel Mareschal, Chairman	http://simu.ulb.ac.be/home.html
CECAM (Centre Européen de Calcul Atomique et Moléculaire)	Instruction and workshops in molecular modeling; funding from 11 research organizations in 9 countries [See site report in App. B.1]	Prof. Michel Mareschal, Director	http://www.cecam.fr/
Austria			
[Multiple sites]	Heterogeneous catalysis		http://majestix.msp.univie.ac.at/agtc_mitglieder.html
Technische Universität Wien: Institute of Physical and Theoretical Chemistry	Electronic structure of solids; magnetism in solids and compounds; *ab initio* molecular dynamics (PAW); parallel computing	Univ. Prof. Dr. Karlheinz Schwarz	http://info.tuwien.ac.at/theochem/; http://www.physchem.tuwien.ac.at/
Technische Universität Wien: Institute of Analytical Chemistry	QSPR methods, neural nets, chemometrics	Prof. Dr. Hans Lohninger	http://qspr03.tuwien.ac.at/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Austria			
Universität Wien: Institut für Materialphysik	Condensed matter physics; heterogeneous catalysis; Wien 97 code; VAMP/VASP (packages for performing *ab initio* molecular dynamics using pseudopotentials and a plane wave basis set)	Univ.-Prof. Dr. Jürgen Hafner; Georg Kresse, Robert Lorenz	http://www.univie.ac.at/materialphysik/; VASP: http://tph.tuwien.ac.at/~vasp/ and http://cms.mpi.univie.ac.at/vasp/vasp/vasp.html
Universität Wien: Institut für Theoretische Chemie und Strahlenchemie	Molecular dynamics and biomolecular simulation	Univ.-Prof. Dr. Othmar Steinhauser	http://www.mdy.univie.ac.at/
	Applied quantum chemistry of molecules, clusters, polymers, metals	Prof. Dr. Alfred Karpfen	http://www.itc.univie.ac.at/~alfred/; http://www.itc.univie.ac.at/qccd.html
	Theoretical biochemistry	Prof. Peter Schuster	http://www.tbi.univie.ac.at/
Belgium			
Solvay	Catalysis; Combinatorial chemistry with Cambridge Combinatorial Ltd. [See Belgium site report in Appendix B.1]	A. Klesing	http://www.cam-com.com/solvay.htm
Janssen Pharmaceutica (part of Johnson & Johnson)	Drug discovery and development, computer-aided prediction of molecular activity [See Belgium site report in Appendix B.1]	Paul Lewi	http://www.janssen-cilag.com/about/r&d/technology2.stm
University of Antwerp, Structural Chemistry Group	Synthesis, characterization and quantum chemical calculations for opto-electronical and sensor devices [See Belgium site report in Appendix B.1]	Prof. Dr. H. J. Geise and C. van Alsenoy	http://sch-www.uia.ac.be/struct/
Katholieke Universiteit Leuven, Chemistry Dept. Quantum Chemistry Group	Gas-phase kinetics and solvent effects, magnetic and optical effects	Dr. Luc Vanquickenborne, Arnout Ceulemans, Minh Tho Nguyen, Marc Hendrickx, Kristin Pierloot	http://www.quantchem.kuleuven.ac.be/research/quantum/quantum_en.html
Université Libre de Bruxelles, Unité de Conformation des Macromolécules Biologiques	Protein modeling	Prof. Shoshona J. Wodak	http://www.ucmb.ulb.ac.be/
Bulgaria			
"Bourgas Prof. Assen Zlatarov University," Lab. of Mathematical Chemistry	Environmental research: QSAR correlations of toxicity with semi-empirical and *ab initio* calculation; reactivity, polymers, molecular topology	Profs. Ovanes Mekenyan and Danail G. Bonchev	http://omega.btu.bg/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Denmark			
Haldor Topsøe	Catalysis research and sales	Dr. Hendrik Topsøe	http://www.haldortopsoe.com/default.htm, http://www.msi.com/pdf/paris98/King_Smith.pdf
Center for Atomic-Scale Materials Physics, Technical University of Denmark	Metallic nanostructures, particularly for catalysis and surface reactivity; DACAPO *ab initio* molecular dynamics code [See site report in App. B.1]	Prof. Jens K. Nørskov	http://www.fysik.dtu.dk/
Finland			
Valtion Teknillinen Tutimuskeskus (VTT, the Technical Research Centre of Finland), Chemical Technology Group	Polymer, catalyst, and lubricant modeling	Dr. Lisbeth Ahjopalo	http://www.vtt.fi/, http://www.vtt.fi/ket/ket4/polymer.htm
Helsinki University of Technology, Laboratory of Physics and Center for Scientific Computing	Visualization, chemistry, physics	Prof. Risto Nieminen	http://www.fyslab.hut.fi/~rni/; http://www.csc.fi/english/
France			
Air Liquide	Designing zeolites for O_2/N_2 separation; complexation of heavy metals [See site report in App. B.1]	Dr. Pluton Pullumbi	http://www.airliquide.com/; http://www.msi.com/solutions/cases/notes/liquide.html
L'Oréal	Cosmetic and dermatological research	Drs. Philippe Barbarat, Roger Rozot	http://www.loreal.com/
Materials Design, s.a.r.l. (Le Mans, France, and Oceanside, CA)	Contract modeling services and MedeA software suite of Windows-based database and search code (InformaticA) and DFT codes [VASP, Augmented Spherical Wave band theory (ElectrA), and linearized muffin-tin orbital (Orestes)] focusing on heterogeneous catalysis inorganic and metallic systems	Dr. Erich Wimmer	http://www.materialsdesign.com/
Rhône-Poulenc (now Aventis)	Surfactants and crystal morphology [Rhône-Poulenc and Hoechst re-organized in 1999 to form Aventis (life sciences), Celanese (Hoechst Chemicals), and Rhodia]	Dr. Marc-Antoine Perrin	http://www.msi.com/solutions/cases/notes/rhone.html

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
France			
Rhône Poulenc / Rhodia	Polymer modeling [Rhodia was formed 1998 from Rhône-Poulenc Chemicals, Fibers and Polymers, absorbing additional components in November 1999] [See site report in App. B.1]	Drs. Caroll Vergelati, Simon Firkins, Theodora Spyriouni	http://www.rhone-poulenc.com/; http://www.pra.org.uk/conf/silicone98.htm (paper 6); http://www.cordis.lu/tmr/src/grants/fmbi/983007.htm
Sanofi-Synthélabo	Pharmaceuticals development	Drs. Jean-Marie Bernassau, Denis Loyaux, Philippe Ochsenbein	http://www.sanofi-synthelabo.fr/us/index.asp
TotalFinaElf	Catalysts, hydrosulfurization, corrosion inhibitors, fuel additives [See site report in App. B.1]	Dr. François Hutschka	http://www.totalfinaelf.com/ho/en/index.htm
GdR 12090: Dynamique Moleculaire Quantique Appliquee a la Catalyse	Joint project: IFP / Totalfina / CNRS / University of Vienna / University of Eindhoven, "*Ab initio* Molecular Dynamics Applied to Catalysis" [See site report in App. B.1]	Director: H. Toulhoat	-
Institut Français du Pétrole	Catalysis and adsorption, supports, chemical additives for fuels and oil production, geochemical problems, polymer properties, thermodynamics and transport properties [See site report in App. B.1]	Dr. Hervé Toulhoat; Dr. Bernard Tavitian; Dr. Philippe Ungerer; Dr. Pascal Raybaud; Dr. Anne-Dominique Biancotto-Defontaines	http://www.ifp.fr/
CEA (Commissariat à l'Énergie Atomique), Le Ripault	Designing optical properties of polymers	Dr. J.P. Dognon	http://paprika.saclay.cea.fr/ang/visite/html/cer_o.htm, http://www.msi.com/solutions/cases/notes/QPA.html
CEA Saclay	Potential model for mixtures of alkanes with multipolar molecules (with Fuchs and Boutin of Univ. Paris-Sud)	Drs. G. Granucci, P. Millié	http://www.saclay.cea.fr/
CEA Grenoble, Dept. for Research on Fundamentals of Condensed Matter (DRFMC)	Condensed matter physics: kinetics Monte Carlo, conventional and *ab initio* molecular dynamics for surfaces and bulk properties	Drs. F. Lançon and T. Deutsch	http://www-drfmc.cea.fr/spm2/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
France			
CNRS / Ecole Supérieure de Chimie Physique Electronique de Lyon (CPE), Laboratoire de Chimie Organométallique de Surface	Heterogeneous and homogeneous catalysis [See site report in App. B.1]	Prof. Jean-Marie Basset, Prof. Philippe Sautet	http://www.cpe.fr/lcoms/basset/default.htm
CNRS - ONERA, Laboratoire d'Etude des Microstructure, Châtillon	Mesoscale modeling of crystal plasticity	Dr. Ladislas P. Kubin	http://zig.onera.fr/lem/
Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble; Laboratoire de Dynamique Moléculaire	Quantum chemistry and molecular dynamics of biological molecules	Dr. Martin Field	http://www.ibs.fr/ext/labos/LDM/welcome_en.htm, http://www.ibs.fr/, http://www.springer-ny.com/catalog/np/may96np/DATA/3-540-60869-9.html
Université de Paris-Sud, Laboratoire de Chimie Physique des Matériaux Amorphes	Modeling thermal diffusion (with B. Hafskjold of Norwegian Univ. of Sci. and Technology; F. Montel of Elf Aquitaine Pau); adsorption in zeolites (with Bernard Tavitian, IFP); molecular potential models (with G. Granucci, P. Millié, CEA Saclay)	Prof. Alain H. Fuchs, Dr. Anne Boutin and Bernard Rousseau	http://www.cpma.u-psud.fr/cpma/cpmahome.html
CNRS / Université Louis Pasteur, Strasbourg	Organometallic catalysis, photochemistry, biological molecules, massively parallel computation (Laboratoire de Chimie Quantique)	Dr. Alain Dedieu	http://www-chimie.u-strasbg.fr/~lcq/
	Methods and techniques of computational chemistry	Dr. Enrico Clementi	
	Theory and modeling of intermolecular interactions (Laboratoire de Chimie Théorique)	Dr. Elise Kochanski	http://www-chimie.u-strasbg.fr/~lct/
	Molecular dynamics, molecular recognition, ions and complexes (Laboratoire de Modélisation et Simulations Moléculaires)	Dr. Georges Wipff	http://www-chimie.u-strasbg.fr/~msm/
Institut Superieur des Materiaux du Mans (ISMANS)	Engineering school (founded 1987) for design and development of materials for technological applications	Dr. François Tsobnang	http://www.ismans.univ-lemans.fr/
ENSIGC-INP, Laboratoire de Génie Chimique	Modeling of gas hydrates	Prof. Jean-Pierre Monfort	

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Germany			
BASF AG	Applications to chemical and materials manufacturing [See site report in App. B.1]	Drs. A. Schäfer, G. Lippert, G. Cox, W. Alenhaufen, J. Sadowski, E. Hädicke	http://www.basf.com/
Bayer AG	Crystallization, process design, catalysis, pharmaceuticals [See Bayer report, App. B.1]	Drs. A. Karbach, R. Bachmann, L. Brüll, F. Reichel	http://www.bayer.com/, http://www.msi.com/solutions/cases/notes/bayer.html
Robert Bosch GmbH	Materials modeling of electronic materials [See report, App. B.1]	Dr. Thomas Brinz	http://www.bosch.de/index_e.htm
COSMOlogic	Software company, spinoff of Bayer to develop and exploit the COSMO-RS methods	Dr. Andreas Klamt	http://www.cosmologic.de
DaimlerChysler, Ulm	Materials research	Dr. Siegfried Doettinger	
Degussa-Hüls (now Degussa)	QSPR for thermophysical properties (from Hüls side before merger); homogeneous catalysis, thermochemistry and thermophysical properties [See site report in App. B.1]	Drs. Robert Franke, Alexander P. Bünz (Marl), Bernd Hannebauer (Wolfgang)	http://www.degussa.de/
Siemens AG	Thermo and kinetics for SiGe epitaxial growth [*J. Vac. Sci Tech. B*, 15, 935 (1997)]	M. Hierlemann, C. Werner and A. Spitzer	http://www.msi.com/materials/cases/CVD2.html; http://tph.tuwien.ac.at/psik/abstracts.html
Forschungszentrum Jülich, Institute of Solid State Research	*Ab initio* calculations for thin magnetic films, nanostructures, and semiconductor surfaces	Dr. Stefan Blügel	http://www.fz-juelich.de/iff/personen/S.Bluegel/
Max-Planck-Institut für Festkörperforschung, Stuttgart	DFT-based *ab initio* molecular dynamics (Car-Parrinello method); properties of solids, liquids, disordered systems and complex molecules	Dr. Michele Parrinello	http://parrserv2.mpi-stuttgart.mpg.de/
	DFT, solids, superconductors, electronic spectra	Dr. Ole Krogh Andersen	http://www.mpi-stuttgart.mpg.de/ANDERSEN/
Max Planck Institute for Polymer Research, Mainz	Polymer modeling	Dr. Kurt Kremer	http://www.mpip-mainz.mpg.de/documents/people/kremer.html
Johannes Gutenberg University of Mainz, Condensed Matter Theory Group	Solids and liquids, glass and other phase transitions	Drs. Rolf Schilling, Kurt Binder	http://www.cond-mat.physik.uni-mainz.de/
Fraunhofer Institute for Manufacturing and Advanced Materials, Bonding Technology and Polymers Dept.	Adhesion of polymers to substrates	Dr. Bernhard Schneider	http://www.ifam.fhg.de/fhg/ifam/frameindex.html?/fhg/ifam/2804/daten_e.html

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Germany			
Fritz-Haber-Institut der Max-Planck-Gesellschaft	Surface and bulk properties of materials	Prof. Dr. Matthias Scheffler	http://www.fhi-berlin.mpg.de/th/th.html
Humboldt University, Berlin	Heterogeneous catalysis	Prof. Dr. Joachim Sauer	http://www.chemie.hu-berlin.de/sonst/agqc/agsauer_eng.html
Universität Augsburg, Institut für Physik	Band theory calculations with augmented spherical wave (ASW) theory and codes	Dr. Volker Eyert	http://www.physik.uni-augsburg.de/~eyert/
Universität Erlangen-Nuremburg, Computer Chemistry Center	*Ab initio* and semi-empirical quantum chemistry; VAMP and TRAMP semi-empirical codes (distributed by Oxford Molecular)	PD Dr. Tim Clark	http://www.ccc.uni-erlangen.de/
Technische Universität Darmstadt	Research and software generation, MOLCAD	Dr. J. Brickmann	http://www.pc.chemie.tu-darmstadt.de/
	Band structure theory; many-body theory	Drs. Jürgen Kübler, N. Grewe	http://www.fkp.physik.tu-darmstadt.de/TheoPhys/
Universität Dortmund	Molecular dynamics simulation (MOSCITO code) and *ab initio* calculations	Prof. Dr. Alfons Geiger	http://ganter.chemie.uni-dortmund.de/
Technical University of Munich	Mechanistic reaction modeling	PD Dr. Werner Thiel	http://forkat.anorg.chemie.tu-muenchen.de/ACI/thiel/wt_thiel_engl.html
Greece			
National Research Centre for Physical Sciences	Solids, polymers, VLE	Prof. Doros Theodorou	http://calypso.nrcps.ariadne-t.gr/
Hungary			
Budapest Univ. of Technology, Inorganic Chem. Dept	Modeling with NMR and crystallography applied to structural problems in materials	Dr. Gábor I. Csonka	http://www.fsz.bme.hu/bme/chemical/csonka.html
University of Veszprém	Organometallic catalysts, Gibbs Ensemble Monte Carlo for VLE	Prof. Lajos Bencze, Dr. Tamás Kristóf	http://mm1.vein.hu/netscape/
Hungarian Academy of Sciences, Research Institute for Solid State Physics and Optics	Metal surfaces, magnetism, muffin-tin orbitals theory; core group member in EU STRUC program	Dr. János H. Kollár	http://www.szfki.hu/, http://www.szfki.hu/~jk/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Italy			
EniChem S.p.A., Centro Ricerche Novara "Istituto Guido Donegani"	Metallocene catalysis [See site report in App. B.1]	Drs. Roberto Fusco, Luca Longo, Paolo Accomazzi, Luigi Abbondanza	http://www.eni.it/; http://www.msi.com/solutions/cases/notes/enichem_full.html
EniTecnologie S.p.A. (jointly owned by Eni, Agip Oils, ENICHEM, Snam and Snamprogetti)	Structure-directing agents for zeolite synthesis; zeolite transport-activity relations; zeolite structure determination [See site report in App. B.1]	Drs. Francesco Frigerio; Roberto Millini; Ricardo Orari	http://www.enitecnologie.it/; http://www.msi.com/solutions/cases/notes/eniricerche_full.html; http://www.msi.com/solutions/cases/notes/anneal.html
Università di Pisa; Dipartimento di Chimica e Chimica Industriale	Molecular modeling of solvation	Prof. Jacopo Tomasi	http://www.dcci.unipi.it/attivita/attivita.html; http://www.dcci.unipi.it/attivita/teorica/dieci.html
Consiglio Nazionale delle Ricerche, Istituto di Chimica Quantistica ed Energetica Molecolare	Chemical and biochemical properties, solutions	Dr. Roberto Ambrosetti	http://www.icqem.pi.cnr.it/text/welcome.html
University of Sassari, Dept. of Agricultural Env. Sciences and Food Biotechnology	Soil chemistry and the biochemistry of plants	Prof. Salvatore Deiana	http://antas.agraria.uniss.it/
SISSA (Scuola Internazionale Superiore di Studi Avanzati)	Theoretical physics - materials	Prof. Erio Tosatti	http://www.sissa.it/cm/
Univ. Trieste; Department of Chemical, Environmental and Raw Materials Engineering (DICAMP)	Molecular simulations for phase equilibrium; rheology of industrial polysaccharides	Drs. Maurizo Fermeglia and Sabrina Pricl	http://www.dicamp.univ.trieste.it/~mau/; http://www.dicamp.univ.trieste.it/personal/sabri/sabri.htm; http://www.dicamp.univ.trieste.it/research/service.asp?page=Rheology
Netherlands			
Akzo Nobel Central Research, Arhem	Polymer fiber modeling and analysis; modeling enantioselectivity (*Journal of Computational Chemistry* 1995, Vol. 16, No. 7, 914)	Dr. Jos Aerts	http://www.msi.com/solutions/cases/notes/PIPD.html
Akzo Nobel / Organon	Pharmaceuticals	Dr. Vicent Gerestein	
DSM Research	Supercritical fluids and thermodynamics, polymers, catalysis, Car-Parrinello first-principles molecular dynamics, molecular vibrational spectra, polymer characterization [See site report in App. B.1]	Prof. Dr. Ludo A.L. Kleintjens and Prof. Dr. Robert J. Meier	http://www.dsm.nl; http://www.dsm.nl/acer; http://www.zam.kfa-juelich.de/compchem/proceedings/Meier/meier.html

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Netherlands			
Philips Research Laboratories, Eindhoven	Partners in EU TMR project on interface magnetism ("*Ab initio* calculations of magnetic properties of surfaces, interfaces and multilayers")[Also see Netherlands report, App. B.1]	Prof. Paul J. Kelly (U. Twente, consultant) and Dr. R. Coehoorn	http://psi-k.dl.ac.uk/TMR1/NL.html
Royal Dutch Shell, Central Research	No activity at this time, but formerly a significant program in molecular simulations and catalysis (Smit, de Swaan Arons, van Santen) [See Netherlands report, App. B.1]		
Holland Research School of Molecular Chemistry	Cooperation between chemistry institutes of the University of Amsterdam (UvA), the Vrije Universiteit (VU) and Leiden University (UL)		http://www.chem.uva.nl/imc/hrsmc/
Netherlands Institute of Catalysis Research (NIOK)	Multi-university graduate program on catalysis: Twente, Amsterdam, Delft, Eindhoven, Leiden, Groningen, Utrecht	Prof. Rutger van Santen, director	http://www.chem.tue.nl/niok/
University of Amsterdam	Molecular simulations of polymers and fluids	Prof. Dr. Ir. Berend Smit	http://molsim.chem.uva.nl/; http://molsim.chem.uva.nl/smit
Technical University of Eindhoven, Schuit Institute of Catalysis	Experimental and theoretical catalysis, involving faculty and students from chemistry, chemical engineering, physics, and mathematics [See site report in App. B.1]	Prof. Rutger van Santen	http://www.tak.chem.tue.nl/ski.html; http://www.tak.chem.tue.nl/people/professors.html
Technical University of Eindhoven, Theoretical Chemistry Group	Heterogeneous catalysis and electrochemical reactions; dynamics Monte Carlo (DMC) simulations; wavepacket dynamics for methane on transition metal surfaces; harm. and anharmonic vibrations	Dr. Tonek Jansen	http://www.tak.chem.tue.nl/theory/
Universiteit Utrecht, Debye Institute	Quantum chemistry methods, radical cations, metal organic chemistry, thermochemistry, and Van der Waals complexes	Prof. Dr. F.B. van Duijneveldt	http://tc5.chem.ruu.nl/Vakgroep.html
Rijksuniversiteit Groningen	DFT and hybrid methods; time-dependent DFT; combined MD and DFT methods; localized electronic states in extended systems	Prof. Dr. J.G. Snijders	http://theochem.chem.rug.nl/
	Biomolecular and aqueous process fluids	Prof. Dr. Herman J.C. Berendsen	http://rugmd0.chem.rug.nl/
U. Leiden	Mesodyn ESPRIT project (international European research on mesoscale modeling)	Dr. Ir. J.G.E.M. Fraaije	http://fwnc7009.leidenuniv.nl/
University of Nijmegen, Institute of Theoretical Chemistry	Intermolecular forces, molecular clusters and collisions, reactive scattering, magnetic properties of larger molecules	Prof. Dr. Ir. Ad van der Avoird	http://www.theochem.kun.nl/; http://quantrs1.theochem.kun.nl/~avda/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Netherlands			
University of Amsterdam		Berend Smit	
Vrije Universiteit	Amsterdam Density Functional code	Prof. Dr. E.J. Baerends	http://www.chem.vu.nl/Staf/baerends/indexEng.html; http://www.ccwp.ac.uk/ccwp/adf.html
FOM Institute for Atomic and Molecular Physics	Molecular simulations for computational physics	Prof. Daan Frenkel	http://www.amolf.nl/; http://www.amolf.nl/external/wwwlab/condensed/compphys/frenkel/
University of Twente, Computational Chemistry Group	Molecular simulations of polymer rheology, sorption and diffusion in polymers and zeolites, reactions in condensed media, crystallization	Prof. W.J. Briels	http://www.ct.utwente.nl/~compchem/
University of Twente, Applied Physics, Computational Materials Science	*Ab initio* calculations of magnetic, optical, electrical, and mechanical properties	Prof. Paul J. Kelly	http://www.tn.utwente.nl/cms/; http://www.tn.utwente.nl/staff/uk/cms.html
Technische Universiteit Delft	Computational physics of materials	Prof. Dr. Simon W. de Leeuw	http://www.cp.tn.tudelft.nl/home/deleeuw/
	Applied thermodynamics and phase equilibrium	Prof. Dr. Ir. J. de Swaan Arons	http://www.stm.tudelft.nl/ttf/ttfhome.htm
Norway			
Norsk Hydro A/S	*Ab initio* calculations on scale formation, zeolites, N_2O decomposition and adsorption, gas hydrates; co-author of GAMESS-UK	Dr. Klaus Schoffel	http://www.hydro.com/; http://www.kjemi.uio.no/~klauss/klaus.html; http://www.uio.no/~vebjornb/hydro/scale/ifp.html
Statoil	Homogeneous catalysis using metallocenes [Thorshaug K., Stavneng, J. A., Rytter, E. and Ystenes M., *Macromolecules* 31 (1998) 7149-7165]	Dr. Erling Rytter	http://www.statoil.no/statoilcom/SVG00990.NSF?opendatabase&lang=en; http://www.msi.com/materials/cases/statoil.html
Norwegian University of Science and Technology, Trondheim, Dept. of Industrial Chemistry	Homogeneous catalysts, structure and vibrational spectroscopy	Prof. Martin Ystenes	http://www.chembio.ntnu.no/~ystenes/
SINTEF Applied Chemistry Institute, Oslo	Homogeneous catalysis	Dr. Ole Swang	http://www.sintef.no/units/chem/catalysis_oslo/molmod.html
Telemark Institute of Technology	Molecular dynamics of water	Dr. Bjørn Kvamme	

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Poland			
Wroclaw University of Technology, Institute of Inorganic Chemistry and Metallurgy of Rare Elements	Inorganic solids and solutions	Prof. Walter Wojciechowski	http://ichn.ch.pwr.wroc.pl/nieorgan/index.xtm
Portugal			
University of Lisbon, Condensed Matter Physics Centre (CFMC)	Simulation of fluid with classical DFT and *ab initio* MD; Telo da Gama on ESF SIMU steering committee	Drs. Margarida Telo da Gama and Benedito J.C. Cabral	http://cfmc.cii.fc.ul.pt/; http://cfmc.cii.fc.ul.pt/Members/BJCC.en.html
Russia			
Moscow State University	Theory; molecular and van der Waals dissociation	N.F. Stepanov	http://www.chem.msu.su/; http://www.chem.msu.su/eng/people/stepanov.html
Institute of Chemical Physics, Moscow			http://www.chph.ras.ru/
Institute of Chemical Physics, Chernogolovka			http://www.icp.ac.ru/
N.D. Zelinsky Institute of Organic Chemistry, Division of Computational Chemistry and Modern Information Technologies	Theory and methods; multibody wave functions; QSAR	N.S. Zefirov	http://www.ioc.ac.ru/; http://www.ioc.ac.ru/Labs/lab44.en.html
Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Moscow			http://www.siobc.ras.ru/
Karpov Institute of Physical Chemistry, Moscow			http://www.nifhi.ac.ru
A.N. Nesmeyanov Institute of Organoelement Compounds, Moscow			http://www.ineos.ac.ru/
Siberian Branch, Russian Academy of Science, Novosibirsk	Catalysis	Boreskov Institute of Catalysis	http://www.catalysis.nsk.su/
	Use of quantum chemistry (Gaussian) for thermochemistry and kinetics of combustion; hydrocarbons, organophosphorous compounds	Institute of Chemical Kinetics and Combustion	http://www-sbras.nsc.ru/eng/sbras/copan/kinetics/
Spain			
Universidad Politecnica de Valencia, Instituto de Tecnologia Quimica (UPV-CSIC)	Heterogeneous catalysis by experimental and computational chemistry methods	Prof. Avelino Corma	http://www.upv.es/itq/; http://www.dupont.com/nacs/Corma.html
Universidad de Sevilla	Simulation of liquid crystals	Prof. Luis F. Rull	http://www.cica.es/aliens/dfamnus/gifelus/rull-ingles.html
Universitat Rovira i Virgili de Tarragona	Molecular simulations for phase equilibrium	Dr. Alan Mackie and Dr. Lourdes Vega	http://www.etse.urv.es/DEQ/web_cat/personal/lvega.htm

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
Sweden			
Lund University, Dept. of Theoretical Chemistry	MOLCAS multiconfigurational *ab initio* code, molecular and solution properties, annual European Summerschool in Quantum Chemistry	Prof. Björn O. Roos	http://www.teokem.lu.se/
Uppsala Universitet, Comp. Chem. Group of Department of Cell and Molecular Biology (ICM)	Q molecular dynamics program	Dr. Johan Åqvist	http://aqvist.bmc.uu.se/main.html
Uppsala Universitet, Institute for Quantum Chemistry (Inst. för Kvantkemi)	Theoretical biochemistry and chemical physics, applied quantum chemistry	Drs. Sten Lunell, Leif A. Eriksson, O. Goscinski	http://falcon.kvac.uu.se/forskning/
Switzerland			
F. Hoffmann-La Roche	Drug discovery [See Hoffmann-La Roche / Novartis site report in App. B.1]	Drs. Frank Grams, Hans-Joachim Bohm, Klaus Mueller	http://www.roche.com/
IBM Zürich Research Laboratory	Development of Car-Parrinello methods and Projector Augmented-Wave method, synthesis and design of novel materials; electronic, structural, dynamical properties of solids and clusters [See site report in App. B.1]	Drs. Wanda Andreoni and Peter E. Blochl	http://www.zurich.ibm.com/Technology/CCP/; http://www.zurich.ibm.com/~and/; http://www.zurich.ibm.com/~blo/
Novartis	Drug discovery [See Hoffmann-La Roche / Novartis site report in App. B.1]	Drs. Romain Wolf and Trevor Payne	http://www.novartis.com/
ETH-Zürich, Dept. of Materials	Stochastic models of polymer dynamics; mesoscale models	Prof. Dr. Hans Christian Ottinger	http://mat.ethz.ch/d-werk/oettinger/welcome.html
	Polymer chemistry and materials modeling	Prof. Ulrich W. Suter	http://mat.ethz.ch/d-werk/suter/
Paul Scherrer Institute	Electronic structure theory and methods (DMol DFT code); modeling zeolites	Dr. Bernard Delley	http://www.psi.ch/; http://www1.psi.ch/www_cmt_hn/delley/
University of Zürich, Institute of Organic Chemistry	Semi-empirical methods (MNDO), coupled quantum mechanics and MD	Prof. Dr. Walter Thiel	http://www.unizh.ch/oci/persons/thiel.html

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
United Kingdom			
Avecia (formerly Zeneca Specialty Chemicals)	Corrosion and scale inhibition [Bromley, L.A., A.M. Buckley, J. Chlad, R.J. Davey, S. Drewe, G.T. Finlan, *J. Colloid Int. Sci.*, 164, 498-502, 1994]	Dr Michael M.H. Charlton	http://www.astrazeneca.com/; http://www.msi.com/solutions/cases/notes/corrosion.html; http://www.msi.com/solutions/cases/notes/scale.html
BNFL	Sorption of gases in zeolites	Dr. Scott L. Owens	http://www.bnfl.co.uk/
BAE (British Aerospace Engineering)	Rare-earth anti-corrosives for aluminum	Dr. Terry Knibb	http://www.bae.co.uk/; http://www.msi.com/solutions/cases/labs/bae/
Advantica Technologies Ltd. (formerly British Gas plc)	GCMC and MD simulation of gas separations using zeolites and polymer membranes; interpreting X-ray diffraction data for minerals; solid-oxide fuel-cell electrolytes [See site report in App. B.1]	Dr. David R. Bates	http://advanticatech.com/; http://www.msi.com/solutions/cases/notes/QPA.html
BP	No UK work at present; past work on development of catalysts and new products, e.g. brake linings, polymerization catalysts, corrosion and scale inhibitors [See RSC site report in App. B.1]	Dr. John Brophy (now at Royal Society of Chemistry), Prof. Nicholas Quirke (now at Imperial College)	
Ciba Specialty Chemicals UK	Pigment stabilization	Dr. Greig Chisholm	http://www.msi.com/solutions/cases/notes/ciba.html
CourtAulds plc	Polymers, biomolecules, rheology	Drs. Jeremy Winter (now at Unilever) and Fiona Case (now at Colgate Palmolive)	
Corus (formerly British Steel)	Modeling organosilanes for corrosion inhibition	Hywel Edwards	http://www.corusgroup.com/
FMC Process Additives Division, Manchester	Additives for scale prevention	Dr. Michael J. Lees	http://www.msi.com/solutions/cases/notes/fmc.html
Fujitsu European Centre for Information Technology Ltd.	Supercomputing applications to materials modeling	Drs. Ross Nobes and Elena V. Akhmatskaya	http://www.fecit.co.uk/Chemistry/
Glaxo Wellcome plc (now GlaxoSmithKline)	Drug discovery, development, and production [See site report in App. B.1]	Dr. Michael M. Hann	http://www.gsk.com/
ICI (Imperial Chemical Industries)	Modeling of nitrocellulose BDEs, polymers; TiO_2 (See Daresbury and RSC reports in App. B.1)	Dr. Andrew Burgess	http://www.ici.com/
Molecular Simulations, Inc. (MSI), subsidiary of Pharmacopeia—British offices	Integrated software for molecular and materials modeling (see site reports for MSI & Pharmacopeia in App. B.3)	Dr. Mike Stapleton	http://www.msi.com/
Oxford Materials Ltd.	Computational materials modeling of paints, coatings and resins, biomaterials, and personal care	Dr. Elizabeth Colbourn	http://www.oxmat.demon.co.uk/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
United Kingdom			
Pfizer Central Research UK	Drug discovery, development, and production	Dr. Catherine Burt	http://www.pfizer.co.uk/recAppInstallDir/c_research/discipline/compchem.htm; http://ala.vsms.nottingham.ac.uk/sbdd/kb.html
Schlumberger Cambridge Research Laboratory	Design cement setting retarders, novel cross-linkers for hydraulic fracturing polymer based fluids, and new scale inhibitors; simulate water swelling of smectite clays to propose new organic clay-swelling inhibitor molecules; study layered double hydroxides and their inorganic and organic intercalates	Dr. Edo Boek; also formerly Dr. Peter Coveney (now at University of London, see below)	http://www.msi.com/solutions/cases/notes/cement.html
SmithKline Beecham plc (now GlaxoSmithKline)	Drug discovery, development, and production [See site report in App. B.1]	Dr. Colin Edge	http://www.gsk.com/
Unilever	Dissolution of powder aggregates under shear flow, deposition of polymers on hair and textiles, transdermal adhesives, pressure-driven flow in nitrocellulose; polymer phase separation; surfactants at oil-water interfaces; Computl. Chemistry Workbench	Drs. Dominic J. Tildesley and Janette Jones	http://www.unilever.com/; http://www.msi.com/solutions/cases/notes/copolymer.html; http://www.msi.com/solutions/cases/notes/oilwater.html; http://hpcc.soton.ac.uk/ProjectDir/app-eng/computational_chemistry.html
Daresbury Laboratory, Theoretical and Computational Science	Research and applications on molecular modeling; Collaborative Computational Projects [See site report in App. B.1]	Dr. Paul Durham	http://wserv1.dl.ac.uk:801/TCSC/groups.html
	Computational Materials Science Group; UKCP	Dr. Nicholas M. Harrison	http://www.cse.clrc.ac.uk/Group/CSECMG/
	Quantum Chemistry Group; GAMESS-UK	Dr. Martyn F. Guest	http://wserv1.dl.ac.uk:801/TCSC/QuantumChem/
	Molecular Simulation Group	Dr. William Smith	http://wserv1.dl.ac.uk:801/TCSC/MolSim/
Royal Society of Chemistry, Molecular Modeling Group	Stimulate the use and application of molecular modeling and related methods in industry [See site report in App. B.1]	Dr. Elizabeth Colbourn, Chair; Dr. Darren R. Flower, Secretary	http://www.rsc.org/lap/rsccom/dab/ind006.htm
UKCP (UK Car-Parrinello Consortium)	Development of Car-Parrinello codes; 9 partners including Daresbury and MSI	Chaired by Prof. Michael J. Gillan, UCL (below)	http://www.cse.clrc.ac.uk/Activity/UKCP/
Consortium for MPP Supercomputing in Materials Chemistry			http://www.hpcc.ecs.soton.ac.uk/hpci/collaborations/materials/about.html
Computational Chemistry Working Party	Collaboration, meetings, shared computing facilities and software	Chaired by N. Handy of Cambridge	http://www.ccwp.ac.uk/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
United Kingdom			
The Royal Institution of Great Britain, Davy-Faraday Research Laboratory	Catalytic sites in titanosilicates and SAPOs	Prof. C. Richard A. Catlow	http://www.ri.ac.uk/DFRL/; http://www.msi.com/solutions/cases/notes/timcm.html; http://www.msi.com/solutions/cases/notes/sapo.html
Edward Jenner Institute for Vaccine Research	Molecular modeling, bioinformatics, and chemoinformatics	Dr. Darren R. Flower	http://www.jenner.ac.uk/
University of Bath	Theoretical condensed matter physics, surface structure, surface chemistry and catalysis, high-energy electron diffraction theory; UKCP	Prof. David Bird	http://www.bath.ac.uk/~pysdb/home.html
Birkbeck College	Crystallographic modeling	Prof. Paul Barnes	
Cambridge University, Department of Chemistry, Theoretical Chemistry	Quantum chemistry, density functional theory (DFT) and theoretical spectroscopy	Prof. Nicholas Handy	http://ket.ch.cam.ac.uk/; http://www.ch.cam.ac.uk/CUCL/staff/nch.html
	Complex liquids or "soft" condensed matter	Prof. Jean-Pierre Hansen	http://www.ch.cam.ac.uk/CUCL/staff/jph.html
	Simulations of liquids and solutions, including hybrid methods	Dr. Michiel Sprik	http://www.ch.cam.ac.uk/CUCL/staff/ms.html
	Potential energy surfaces for transition states and glass structures	Dr. David Wales	http://www.ch.cam.ac.uk/CUCL/staff/djw.html
Cambridge University, Department of Physics, Theory of Condensed Matter Group	CASTEP (Cambridge Serial Total Energy Package); CETEP (Cambridge Edinburgh Total Energy Package) for massively parallel computers; UKCP	Prof. Michael C. Payne	http://www.tcm.phy.cam.ac.uk/
Cambridge University, Department of Materials Science and Metallurgy	Modeling interfaces and defects in solid-state materials, particularly grain boundaries; UKCP	Dr. Paul D. Bristowe	http://www.msm.cam.ac.uk/asg/
University of Edinburgh, Department of Chemical Engineering	Modeling of adsorption	Prof. Nigel Seaton	http://lyne.chemeng.ed.ac.uk/research/Seaton.html; http://www.chemeng.ed.ac.uk/people/nigel.html
University of Edinburgh, Dept. of Physics	Computational condensed matter theory; UKCP	Dr. Graeme Ackland	http://www.ed.ac.uk/~graeme/
Imperial College of Science, Technology and Medicine; Department of Chemistry	New methods and codes for molecular simulations, applied to problems in materials, catalysis, nanotechnology, cheminformatics and synthesis [See RSC site report in App. B.1]	Prof. Nicholas Quirke and Dr. Henry Rzepa	http://www.ch.ic.ac.uk/quirke/; http://www.ch.ic.ac.uk/rzepa/; http://www.ch.ic.ac.uk/computational/; http://www.ch.ic.ac.uk/research.html
Imperial College of Science, Technology and Medicine; Dept. of Chem. Eng.	Monte Carlo simulation of liquid crystals; equations of state [See RSC site report in App. B.1]	Dr. George Jackson	http://www.ce.ic.ac.uk/mfs/gj.htm

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
United Kingdom			
University of Kent, Center for Materials Research	Molecular simulations to interpret NMR and EXAFS	Dr. Alan V. Chadwick	
University of Leeds, Interdisciplinary Research Centre in Polymer Science and Technology	Modeling of polymers	Prof. T.C.B. McLeish	http://www.irc.leeds.ac.uk/irc/
University of London, Queen Mary & Westfield College	Materials modeling by molecular simulation, particularly mesoscale modeling of soft condensed matter	Dr. Peter Coveney	http://www.chem.qmw.ac.uk/research/pvc.html
University College London, Department of Physics and Astronomy	Surface chemistry, liquid semiconductors; Conquest O(N) code; UKCP chair	Prof. Michael J Gillan	http://www.keele.ac.uk/depts/ph/research/tc/cph_mem/mjg.html; http://www.cmmp.ucl.ac.uk/~conquest/
University College London, Department of Geological Sciences	*Ab initio* predictions for mineral and crystal structure and properties	Prof. G. David Price	http://www.ucl.ac.uk/GeolSci/people/d-price/gdp.htm
University College London, Biomolecular Structure and Modeling	Analyze and predict protein structure and function	Prof. Janet M. Thornton	http://www.biochem.ucl.ac.uk/bsm/biocomp/; http://www.biochem.ucl.ac.uk/~thornton/
University of Manchester, Institute of Science and Technology, Molecular Simulation Laboratory	Organic and biological modeling	Prof. Julian Clarke	http://molsim.ch.umist.ac.uk/
Oxford Centre for Molecular Sciences	Molecular simulations	Dr. W.G. Richards	
Oxford University, Phys. & Theoretical Chemistry Laboratory	Molecular simulation of zeolites, glasses; UKCP	Prof. Paul A. Madden	http://www.chem.ox.ac.uk/researchguide/pamadden.html
Oxford University, Theoretical Physics, Condensed Matter Group	Soft condensed matter: phase equilibria and rheology for miscible and immiscible binary and ternary fluids, amphiphilic systems, polymer solutions and colloidal suspensions	Dr Julia M. Yeomans	http://www-thphys.physics.ox.ac.uk/users/JuliaYeomans/
Oxford University, Department of Earth Sciences	Model dissociation of water and hydration and detachment of cations for natural weathering reactions of rocks and the trapping and transport of man-made toxic metals by natural mineral systems (*Phys Rev. B* 52, 1995, 10823); MolDy code; UKCP	Dr. Keith Refson	http://www.earth.ox.ac.uk/~keith/

Table 10.1 (continued)
Recent and Ongoing Modeling Activities in Europe

Location	Type of activity	Contact(s)	Web links
United Kingdom			
Oxford University, Department of Materials, Materials Modeling Laboratory	Modeling of materials across length scales from electronic to continuum	Prof. David G. Pettifor	http://www.materials.ox.ac.uk/MML/
Queen's University of Belfast, School of Mathematics and Physics, Atomistic Simulation Group	Nanoscale materials modeling; UKCP	Prof. Michael W. Finnis	http://titus.phy.qub.ac.uk/
University of Surrey, Dept. of Chemistry	Molecular simulations	Dr. David Heyes	
University of Warwick, Department of Chemistry	Inhibition of clathrate hydrate formation, chiral nanoporous materials	Dr. P. Mark Rodger	http://www.warwick.ac.uk/fac/sci/Chemistry/astaff/pmrodger/

Table 10.2
Recent and Ongoing Modeling Activities in Japan

Location	Type of activity	Contact(s)	Web links
Asahi Chemical Industry Co., Ltd	Metallocene catalysis, polymer properties including mesoscale modeling [See site report in App. B.2]	Dr. Tomosumi Yamada	http://www.asahi-kasei.co.jp/asahi/english/
Banyu Pharmaceutical Co., Ltd.	Pharmaceutical discovery and development	Mr. Masaya Seko	
Eisai Co., Ltd.	Pharmaceutical discovery and development; key role in developing Aricept, an Alzheimer's disease medication [See site report in App. B.2]	Dr. Yohiyuki Kawakami	http://www.eisai.co.jp/; http://www.aricept.com/
Fuji Photo Film, Ashigara Research Laboratory	Polymers; part of Doi Project [See description in Short Reports, App. B.2]	Dr. Hiroo Fukunaga	http://home.fujifilm.com/
Fuji Xerox Co., Ltd.	Development of special-purpose MD computer with Taisho Pharmaceuticals, Tottori Univ., and Nagoya Univ.; force fields and MD of amorphous PTFE (*Mol. Sim.*, 1999, 21, pp. 325-342)	Okimasa Okimada	
Fujisawa Pharmaceutical Co., Ltd. (Osaka)	Pharmaceutical discovery and development	Dr. Nakanishi	
Fujitsu Ltd.	Software development and services: MOPAC2000, Mozyme, MASPHYC, CAChe packages [See site report in App. B.2]	Munetaka Takeuchi, Masahito Yamaguchi, Takashi Mitsui	http://fujitsu.co.jp/; http://www.cache-molecular.com/
Fujitsu Laboratories Ltd.	Developed MOS-F (predicts spectroscopic properties of large organic systems); find electronic and optical materials efficiently [See site report in App. B.2]	Toshiaki Narusawa, Azuma Matsuura	http://fujitsu.co.jp/

Table 10.2 (continued)
Recent and Ongoing Modeling Activities in Japan

Location	Type of activity	Contact(s)	Web links
Japan Polyolefins, Ltd.	Polyethylene properties and synthesis, including metallocene catalysis; Doi Project [See description in Short Reports, App. B.2]	Hiroyasu Tasaki	http://www.sdk.co.jp/
Japan Tobacco Inc. (Osaka)	Pharmaceutical discovery and development	Dr. Okada, Dr. Okajima	
JSR Corporation	Metallocene synthesis and properties of polymers [See description in Short Reports, App. B.2]		http://www.jsr.co.jp/
Kao	Consumer products	Dr. Tadahiro Ozawa	
Mitsubishi Chemical Corp.	Design and property-prediction for chemicals and materials; metallocene catalysts, polymer membranes, dyes, surfactants, dispersants [See site report in App. B.2]	Mr. Yukikazu Natori, Dr. Takeshi Ishikawa	http://www.m-kagaku.co.jp/index_en.htm
Mitsubishi Heavy Industry Co.	Materials modeling; part of CAMP project	Dr. Izumi	
Mitsubishi-Tokyo Pharmaceuticals Inc.	Pharmaceutical discovery and development; part of JBIC (Japan Bioindustry Consortium) [See site report in App. B.2]	Drs. Hiroaki Ueno and Hideo Kubodera	http://mitsubishi-pharm.co.jp/
Mitsui Chemical Inc. [Merging with Sumitomo Chemical, 2000-2003]	Polymers, catalyst design, materials development, biochemicals, and process design; part of Doi Project [See description in Short Reports, App. B.2]	Drs. Heng Phara, Kunio Sannohe, Takashi Igarashi	http://www.mitsui.co.jp/
NEC, Fundamental Research Laboratories	Development of codes for NEC computers (AMOSS, VRMS) and of special-purpose computers; applications to modeling semiconductors, insulators, processing, environmental concerns; part of CAMP project [See site report in App. B.2]	Drs. Hitoshi Igarashi, Akihiro Iwaya, Toshikazu Takada, Yoshiyuki Miyamoto, Shuhei Ohnishi	http://www.nec-global.com/
Nippon Zeon Co., Ltd.	Polymer properties and catalytic synthesis; part of Doi Project [See description in Short Reports, App. B.2]		http://www.zeon.co.jp/
Sankyo Co., Ltd.	Chemicals and pharmaceuticals	Dr. Shuichi Miyamoto	

Table 10.2 (continued)
Recent and Ongoing Modeling Activities in Japan

Location	Type of activity	Contact(s)	Web links
Sumitomo Chemical Co., Ltd. [Merging with Mitsui Chemicals, 2000-2003]	Chemicals, life sciences, and materials: agro-chemicals (pesticide design), functional dyes, polymers (including liquid crystals), catalysts (Ziegler-type), photo-electric materials, and synthesis routes for pharmaceuticals; part of Doi Project, CAMP project [See description in Short Reports, App. B.2]	Drs. Ishida, Zempo, Katsuhiro Suenobu	http://www.sumitomo-chem.co.jp/
Taisho Pharmaceutical Co., Ltd.	Pharmaceutical discovery and development, MD modeling of membrane, special hardware for MD (MD engine) [See site report in App. B.2]	Dr. Kunihiro Kitamura	http://www.taisho.co.jp/
Takeda Chemical Industries, Ltd.	Pharmaceutical discovery and development [See site report in App. B.2]	Dr. Yoshio Yamamoto	http://www.takeda.co.jp/
Tanabe Pharmaceutical Co., Ltd.	Pharmaceutical discovery and development	Dr. Hiroki Shirai	
Toshiba R&D	Silicon oxidation, dendrimers, quantum chemistry and molecular simulation in reaction-engineering tools and materials design; part of CAMP project [See site report in App. B.2]	Yasuo Ikawa, Drs. Koichi Kato, Shigenori Tanaka, Satoshi Itoh, Shinji Onga, Yuusuke Sato	http://www.toshiba.com/
Toyota Central Research & Development Laboratories	Materials modeling for high-temperature and high-power electronic devices; modeling for environmental issues; Part of CAMP project [See site report in App. B.2]	Dr. Atsuo Fukumoto	http://www.global.toyota.com/
UBE Industries, Ltd.	Polymers and polymerization catalysts; part of Doi Project [See description in Short Reports, App. B.2]	Dr. Koga, Tatsuya Shoji	http://www.ube-ind.co.jp/
Yamanouchi Pharmaceutical Co., Ltd.	Pharmaceutical discovery and development; solution of X-ray structures	Dr. Fujita	http://www.yamanouchi.co.jp/
Japan Chemistry Program Exchange	Distribution of computational chemistry codes (analogous to QCPE in U.S.)		http://www.aist.go.jp/NIMC/
CAMP and CAMM projects	Multi-industry project on materials research [See NEC site report in App. B.2]	Dr. Shuhei Ohnishi (coordinator)	http://www.camp.or.jp/
"Platform for Designing High Functional Materials," the Doi Project	MITI/NEDO-sponsored project on multiscale modeling of polymers involving 11 companies and 5 universities [See site report in App. B.2]	Dr. Masao Doi	http://www.stat.cse.nagoya-u.ac.jp/

Table 10.2 (continued)
Recent and Ongoing Modeling Activities in Japan

Location	Type of activity	Contact(s)	Web links
AIST (National Institute of Advanced Industrial Science and Technology)	Molecular dynamics simulation of peptide:DNA complex (*Mol. Sim.*, 1999, 21, pp. 303-324) using PEACH software (*J. Comput. Chem.*, 1997, 18, 1546) on GRAPE-MD processor [See JRCAT site report in App. B.2]	Dr. Yuto Komeiji	http://staff.aist.go.jp/y-komeiji/
Institute for Molecular Science (IMS), Okazaki	National supercomputer center for molecular science community; modeling protein folding, solution chemistry, organic solids, excited-state molecules, spectroscopy; sponsors Quantum Chemistry Literature Database (QCLDB) [See site report in App. B.2]	Drs. Suehiro Iwata, Hiroki Nakamura, Fumio Hirata, Yuko Okamoto	http://www.ims.ac.jp; http://qcldb.ims.ac.jp
Institute for Solid State Physics (ISSP)	National supercomputer center for condensed-matter physics community; research on prediction and design of material properties, including electronic materials [See site report in App. B.2]	Professors Hidetoshi Fukuyama and Hajime Takayama	http://www.issp.u-tokyo.ac.jp
Joint Research Center for Atom Technology (JRCAT), Theory Group, Tsukuba	Electronic structure of solids; atomic and molecular processes in solids and on solid surfaces [See site report in App. B.2]	Dr. Kiyoyuki Terakura	http://www.jrcat.or.jp/
NIRIN, National Industrial Research Institute of Nagoya	Dynamics at ceramic grain boundaries, design by computer simulation, mechanical deformation of ZrO_2 polycrystal	Dr. Hiroshi Ogawa	http://www.nirin.go.jp/People/ogawa/
National Institute for Fusion Science	Cluster structure formation in short-chain polymers by molecular dynamics	Dr. Susumu Fujiwara	http://www.tcsc.nifs.ac.jp/; http://www.tcsc.nifs.ac.jp/fujiwara/
RIKEN, The Institute of Physical and Chemical Research	Special-purpose computers for molecular dynamics, homology search; interpretation of X-ray diffraction for 3-D structure of matter [See site report in App. B.2]	Dr. Toshikazu Ebisuzaki	http://www.riken.go.jp/eng/; http://www.riken.go.jp/eng/Yoran/RA/csl.html
Tohoku National Industrial Research Institute	MD and first-principles MD for clusters, water, grain boundary; modeling of supercritical water	Drs. Tamio Ikeshoji and Chee Chin Liew [See JRCAT site report in App. B.2]	http://www.tniri.go.jp/outline/section/kiso/ikeshouji-e.html; http://www.aist.go.jp/TNIRI/~liew/
Chiba University, Department of Chemistry	Modeling of adsorption	Prof. Katsumi Kaneko	http://pchem2.s.chiba-u.ac.jp/eng/
Hosei University	Molecular simulation for diffusion near the critical point	Prof. Yosuke Kataoka	
University of Hiroshima	*Ab initio* molecular dynamics of reactions	Prof. Misako Aida	http://www.chem.sci.hiroshima-u.ac.jp/people/index.html
	Computational Chemistry for Atmospheric Environmental Molecules (CCAEM)	Prof. Suehiro Iwata (also at IMS)	http://hera.chem.sci.hiroshima-u.ac.jp/ccaem/index.html

Table 10.2 (continued)
Recent and Ongoing Modeling Activities in Japan

Location	Type of activity	Contact(s)	Web links
Kanazawa University, Department of Computational Science	First-principles molecular dynamics of materials, polymer modeling with special-purpose computers	Prof. Yasuaki Hiwatari (Chair, Molecular Simulation Society of Japan)	http://cmpsci.s.kanazawa-u.ac.jp/~actrep/98/act.hiwa.html; http://www.echem.titech.ac.jp/mssj/ (In Japanese)
Keio University	Molecular simulations	Prof. S. Nosé	
Kochi University	Molecular dynamics for liquid-solid mixtures of salts (with K. Kawamura of Earth and Planetary Sciences, Tokyo Institute of Technology; *Mol. Sim.*, 1999, 21, pp. 387-399)	Tadashi Akamatsu	
Kurashiki University of Science and Art, Dept. of Chemistry	Molecular simulations for vapor-liquid equilibrium	Prof. Koichiro Nakanishi	http://www.kusa.ac.jp/%7Enakanisi/
Kyoto University	Protein modeling	Prof. Nobuhiro Go	http://www.qchem.kuchem.kyoto-u.ac.jp/
	Quantum physical chemistry	Prof. H. Nakatsuji	http://quanta.synchem.kyoto-u.ac.jp/ (in Japanese)
Meiji University	Adsorption on zeolites (expt and modeling)	Dr. K. Chihara	http://www.meiji.ac.jp/servers/chem/faculty-e.html
Osaka University, Dept. of Chemical Science and Engineering	Molecular simulation for supercritical fluids, adsorption	Prof. Tomoshige Nitta, Kazunari Ohgaki	http://lab4-7.cheng.es.osaka-u.ac.jp/nittalabo/nittalabo-e.html
Osaka University, Institute for Protein Research	Protein modeling	Prof. Haruki Nakamura	http://www.protein.osaka-u.ac.jp/kessho/members/harukin/index-eng.html
University of Nagoya, Dept. of Applied Physics Polymers	Biomolecules, rheology [See "Doi Project" site report in App. B.2]	Prof. Masao Doi	http://www.stat.cse.nagoya-u.ac.jp/
University of Nagoya, Department of Chemistry	Water and supercritical fluids [See site report in App. B.2]	Prof. Iwao Ohmine	http://www.chem.nagoya-u.ac.jp/ohmineken/omineken_home_Eng.html
Tohoku University, Institute for Chemical Reaction Science	Modeling of reactions	Prof. Nobuyuki Harada	http://www.icrs.tohoku.ac.jp/research/info/main-e.html
Tohoku University, Laboratory for Molecular Materials Design	Catalysts, membranes, ceramics, electronic devices, tribology; integrating computational chemistry methods	Prof. Akira Miyamoto	http://www.che.tohoku.ac.jp/labo/Miyamoto/
Tohoku University, Dept. of Engineering Science	Molecular simulations	Dr. M. Tanaka	http://www.che.tohoku.ac.jp/labo/Tanaka/index.html
Tokyo Institute of Technology	Lipid membranes, aqueous salt solution, parallel molecular dynamics	Prof. Susumu Okazaki	http://www.echem.titech.ac.jp/okazaki/okazaki.html
Univ. of Tokyo, Applied Chemistry	Multi-configuration methods, applications to material design: nanostructures, chiral crystallization, resolving reagents	Dr. Kimihiko Hirao	http://www.appchem.t.u-tokyo.ac.jp/appchem/labs/hirao/index-e.html

Table 10.3
Modeling Activities in Africa, Australia, and Asia Outside Japan

Location	Type of activity	Contact(s)	Web links
Australia			
Australian National University, Research School of Chemistry	Methods of computational quantum chemistry, computational organic chemistry, free radical chemistry, enzyme-mediated chemistry, polymer chemistry	Dr. Leo Radom	http://rsc.anu.edu.au/RSC/ChemResearch/Chemists/radom.html; http://rsc.anu.edu.au/~radom/
	Molecular simulations	Dr. Dennis Evans	http://rsc.anu.edu.au/RSC/ChemResearch/Chemists/evans.html
Brisbane Molecular Modeling Group	Academic research on computational chemistry applications including drug design	Dr. Michael Dooley	http://www.uq.edu.au/~ddmdoole/
Computational Chemistry Project; ACCVIP (Australian "Computational Chemistry via the Internet" Project)	Educational modules and online courses	Drs. E.J. Lloyd (Monash Univ.), B. Salter-Duke (Northern Territory Univ.), M.G. Wong (Swinburne Univ. of Technology), B.F. Yates (Univ. Tasmania), G.J. Cross (Monash Univ.)	http://www.chem.swin.edu.au/CAUT.html
Monash University, Dept. of Medicinal Chemistry	Drug design by molecular modeling, QSARs and neural networks	Dr. Edward J. Lloyd	http://www.chem.swin.edu.au/members/e_lloyd.html
University of Tasmania, Catalysis and Catalyst Modelling Group	Biologically active materials, reaction mechanisms of organic chemistry	Dr. Brian F. Yates	http://www.chem.utas.edu.au/staff/yatesb/
Hong Kong			
Hong Kong University	Molecular simulations for materials properties tied to data	Dr. K.Y. Chan	http://chem.hku.hk/~kyc/kychan.html
Hong Kong University of Science and Technology	Quantum chemistry for catalysis and materials modeling	Dr. Zhenyang Lin	http://home.ust.hk/~chzlin
India			
Indian Institute of Science, Bangalore; Inorganic and Physical Chemistry	Polymers, chemisorption and catalysis, nanotechnology, electrochemistry, biophysical chemistry	Drs. K.L. Sebastian and B.J. Cherayil	http://ipc.iisc.ernet.in/
Indian Institute of Science, Bangalore; Organic Chemistry	Molecular simulations and modeling for organic reactivity, magnetic and nonlinear optical properties	Dr. Jayaraman Chandrasekhar	http://www.orgchem.iisc.ernet.in/faculty/
National Centre for Biological Sciences, Bangalore	Protein folding and unfolding; code development	Dr. R. Sowdhamini	http://www.ncbs.res.in/%7Efaculty/mini.html
National Chemical Laboratory, Pune	Quantum chemistry for catalysis: zeolites and organometallics; spectroscopy	Drs. S. Pal, S.D. Prasad, R. Tewari, R. Vetrivel	http://www.ncl-india.org/research/theory/theory.html

Table 10.3 (continued)
Modeling Activities in Africa, Australia, and Asia Outside Japan

Location	Type of activity	Contact(s)	Web links
Israel			
Lise Meitner-Minerva Center for Computational Quantum Chemistry	Catalysis and bio-organic molecules	Profs. Sason Shaik (Hebrew Univ. of Jerusalem) and Yitzhak Apeloig (Technion)	http://alpha.ch.huji.ac.il/public_html/index.htm
Technion, Institute for Advanced Studies in Theoretical Chemistry (IASTC)	Workshops, short courses, code development	Prof. Nimrod Moiseyev	http://www.technion.ac.il/technion/chemistry/IASTC/
Tel Aviv University, Theoretical Chemistry	Theory and modeling of materials and reactions		http://www.tau.ac.il/chemistry/faculty/theochem.html
Weizmann Institute of Science, Department of Chemical Physics	Quantum kinetics, molecular dynamics of diffusion and reaction in the liquid phase	Prof. Anatoly Burshtein	http://www.weizmann.ac.il/chemphys/MD.shtml
Kenya			
University of Nairobi, Department of Chemistry	Quantum chemistry and spectroscopy	Prof. David Njoroge Kariuki	http://www.uonbi.ac.ke/acad_depts/chemistry/dnkariuki.html
Korea			
KAIST	Computer simulation including molecular dynamics, Monte Carlo methods and molecular modeling on biomolecules in water environment	Prof. Mu-Shik Jhon	http://muse.kaist.ac.kr/
Keimyung University	Molecular simulations	Prof. S.-H. Suh	
New Zealand			
University of Canterbury	Quantum chemistry for molecular properties including transport properties	Dr. Robert G.A.R. Maclagan	http://www.chem.canterbury.ac.nz/people/academics/rgarm.htm
Pakistan			
Husein Ebrahim Jamal Research Institute of Chemistry	Quantum chemistry for modeling for naturally occurring bioactive substances	Dr. Muhammad Iqbal Choudhary	http://www.hejric.edu/; http://www.hejric.edu/iqbal.html
Saudi Arabia			
King Fahd University of Petroleum and Materials	Modeling of spectroscopy, polymers, biological compounds, liquid crystals	Dr. Mohammed A. Morsy, Hassan M. Badawi, Wolfgang Forner	http://www.kfupm.edu.sa/chem/
Taiwan			
National Chiao Tung University, Institute of Applied Chemistry	Self-assembled monolayers on gold	Dr. Ta-Wei Li	http://www.nctu.edu.tw/english/colleges/cs/dac.html
Institute of Biomedical Sciences, Academia Sinica	Life sciences	Dr. Carmay Lim	http://ibms.sinica.edu.tw:8080/~carmay/c1.html

Table 10.3 (continued)
Modeling Activities in Africa, Australia, and Asia Outside Japan

Location	Type of activity	Contact(s)	Web links
South Africa			
University of Cape Town	Polymer modeling, free energy calculation methods, solvation structure, NMR, organometallic catalysis, biopolymers	Dr. Kevin J. Naidoo	http://hydrogen.cem.uct.ac.za/compchem/
University of the North, Materials Modeling Centre	Molecular simulations	Phuti E. Ngoepe	
University of the Witwatersrand, Molecular Science Institute	Modeling of electrochemistry, crystallography, drug and molecule design	Profs. T. Koritsanszky, D.C. Levendis, F. Marsicano, L. Glassser, I. Cukrowski	http://www.chem.wits.ac.za/

Table 10.4
Recent and Ongoing Modeling Activities in North and South America Outside the United States

Location	Type of activity	Contact(s)	Web links
Brazil			
Rio de Janeiro Federal University (UFRJ), Institute of Chemistry	Modeling atmospheric chemistry, spectroscopy, adsorption	Prof. Marco Antonio Chaer Nascimento	http://www.iq.ufrj.br/~chaer/
University of São Paulo, Institute of Physics	Magnetic properties: layered Fe and Cr using Linear Muffin-Tin Orbitals; impurities in Au	Prof. Sonia Frota-Pessôa	http://www.sbf.if.usp.br/WWW_pages/Eventos/xxenfmc/res0698.ps and res0699.ps
Canada			
Uniroyal Chemical Ltd.	Design and synthesis of insecticides	Dr. Mark A. Dekeyser	http://www.uniroyal.ca/
Xerox (Canada)	Modify and develop photocopying materials, crystals, polymers (See *I&EC Research* 34, 4174-4184, 1995)	Dr. Thomas A. Kavassalis	http://www.xerox.com/
NRC Steacie Institute for Molecular Sciences, Theory and Computation Program	Molecular spectroscopy, photomanipulation, thermochemistry, solids, interfaces, biochemical mechanisms	Dr. John Tse, Prof. Dennis R. Salahub	http://www.sims.nrc.ca/sims/theory_e.html
Université de Montréal, Centre de Recherche en Calcul Appliqué (Centre for Research on Computation and its Applications – CERCA) and Department of Chemistry	DFT of transition metals; surfaces, clusters, chemisorption, dynamic surface phenomena, materials, pharmaceutical chemistry	Prof. Dennis R. Salahub	http://www.cerca.umontreal.ca/profs/salahub.html; http://www.fas.umontreal.ca/CHIMIE/employes/SALAHUBDennis.html
University of Calgary	Homogeneous catalysis, density functional theory	Prof. Tom Ziegler	http://www.cobalt.chem.ucalgary.ca/group/master.html

Table 10.4 (continued)
Recent and Ongoing Modeling Activities in North and South America Outside the United States

Location	Type of activity	Contact(s)	Web links
Canada			
Queen's University, Kingston	Development of density functional theory methods, basis-set-free MO calculations	Prof. Axel D. Becke	http://www.chem.queensu.ca/faculty/becke/beckeh.html
McMaster University	Atoms in Molecules method	Prof. Richard F.W. Bader	http://www.chemistry.mcmaster.ca/faculty/bader/bader.html
Mexico			
Instituto Mexicano del Petróleo, Institutional Prog. on Molecular Simulation	Molecular simulations R&D for the Mexican Petroleum Corporation (PEMEX)	Drs. Marcelo Lozada y Cassou and Beatriz Castro	http://www.imp.mx/
Instituto Fisica, UNAM	Free energy of solvation for the dissolution of organic compounds in water (with Mackie and Vega, U. Rovira i Virgili, Spain)	Dr. Jorge Hernandez-Cobos	
National Univ. Mexico; Instituto de Quimica, Physical Chemistry	Predicting ligand-receptor interactions in molecularly imprinted polymers used for the removal of organosulfur compounds from fuels	Dr. Raul Cetina	http://www.iquimica.unam.mx/
Venezuela			
Intevep (affiliate of Petróleos de Venezuela)	MD simulations of surfactants at oil/water interfaces and of asphaltenes	Drs. Maria Lupe Marquez and Estrella Rogel	http://www.intevep.pdv.com/ingles/
Instituto Venezolano de Investigaciones Científicas (IVIC), Centro de Química	Laboratorio de Química Computacional: Catalysis, effective Hamiltonians and effective potentials	Drs. Fernando Ruette and Aníbal Sierraalta	http://bragg.ivic.ve/Ivic/IvicCQ/labs/quim.comp/comput.html (in Spanish)
	Laboratorio de Espectroscopia Molecular: Adsorption on metal oxides (with Bader), molecular modeling of individual asphaltene molecules (Murgich et al. *Energy & Fuels* 1996, 10, 68)	Dr. Juan Murgich	http://quimica.ivic.ve/quimica/jmurgich.cvspan.html
Universidad Simón Bolívar	Application of molecular simulation to problems in the oil industry: adsorption/phase equilibria/asphaltenes/Joule-Thomson coefficient	Dr. Erich A. Müller	http://www.tadip.tf.usb.ve/emuller/molsim.htm

Table 10.5
Recent and Ongoing Commercial Modeling Activities in the United States

Location	Type of activity	Contact(s)	Web links and references
3M, Inc.	Design of hydrofluoroethers, fluorescent potassium monitor, additives for imaging films, fluorescent pigment, dental polymers; modeling of polymer structure, binding, and spectroscopy; catalysts [See site report in App. B.3]	Dr. Cristina U. Thomas	http://www.mmm.com/
Air Products and Chemicals, Inc.	Structure-property calculations for selective fluorination and nitration, thermochemistry, chemical vapor deposition, homogeneous catalysis, gas adsorption and storage, metal dusting, chemical corrosion, heterogeneous catalysis [See site report in App. B.3]	Drs. Brian K. Peterson and Hansong Cheng	http://www.airproducts.com/
Albemarle Chemicals	Flame-retardant chemicals (e.g., hexabromo-cyclododecane); methyl aluminoxane models (Simeral GOALI project at LSU with Randall Hall and NIST ATP project)	Drs. Gary Zhao and Larry S. Simeral	http://www.albemarle.com/; http://www.nist.gov/public_affairs/atp2000/00004278.htm
Alcoa	Vinylphosphonic acid coating on Al for adhesion; VASP and PW91 to develop coating for roller-mill abrasion protection	Drs. Louis G. Hector, Jr. (now at General Motors Research Laboratory), G. Nitowski, and A. Maiti	http://www.alcoa.com/; http://www.msi.com/materials/cases/coating.html
American Cyanamid (now part of American Home Products)	QSAR development and design of insecticidal uncouplers based on pyrrolomycin and related pyrrole insecticides	Dr. David M. Gange (now of Intercardia Research Labs)	http://lab-robotics.org/Mid_Atlantic/meetings/bio_gange.htm
ARCO Chemicals (now part of Lyondell)	Modeling of polymer weatherability and design of improvements		
Bayer, Bushy Park SC	Modeling of pigment crystals [See site report in App. B.1]	Dr. Gamil Guirgis	http://www.bayer.com/
B.F. Goodrich	Poly(norbornene) development, now discontinued	Prof. Peter Ludovice, GA Tech	
BP (formerly Amoco), Naperville, IL	Heterogeneous and homogeneous catalysis, thermochemistry, product and process development, crystal structures	Dr. Joseph T. Golab	http://www.bp.com/
Bristol-Myers Squibb	Composition from Rietveld analysis of crystal forms in a two-component mixture	M. Davidovich, R.L. Mueller, I.M. Vitez	http://www.msi.com/materials/cases/bms-rietveld.html

Table 10.5 (continued)
Recent and Ongoing Commercial Modeling Activities in the United States

Location	Type of activity	Contact(s)	Web links and references
Calgon Inc. (now part of Nalco / Suez Lyonnaise des Eaux)	QSAR analysis for scale-inhibitor design	Dr. Beverley Bendiksen	http://www.nalco.com/; http://www.msi.com/materials/cases/qsar1.html
CambridgeSoft.com	ChemDraw, Chem3D, ChemFinder, ChemOffice software for molecular modeling and chemical data management [See site report in App. B.3]	Dr. Bruce R. Gelin	http://www.cambridgesoft.com/; http://www.chemfinder.com/; http://www.chemnews.com/
Chamot Laboratories, Inc.	Computational chemistry contract services and software training	Dr. Ernest Chamot	http://www.chamotlabs.com/
Chevron Corporation, Petroleum Energy and Environmental Research Lab, Caltech; merging into ChevronTexaco	Oilfield applications: corrosion, lubrication, wear, catalysis, geochemistry	Dr. Yongchun Tang	http://www.chevron.com/; http://www.peer.caltech.edu/tang.htm
Colgate Palmolive	Predicting surfactant properties and mesoscale modeling	Dr. Fiona H. Case	http://www.colpal.com/
Delphi Research Laboratories (1999 spinoff from General Motors)	Apply density functional theory to study adhesion of Nb, Al, Ag with Al_2O_3 [*Phys. Rev. Letters*, 85:15, 3225-3228 (2000)]; impurity effects of adhesions of Fe and Al	Dr. John R. Smith	http://www.delphiauto.com/dti/; http://publish.aps.org/abstract/PRL/v85/p3225
The Dow Chemical Company	Low-dielectric polymers, homogeneous and heterogeneous catalysis, thermochemistry and physical properties [See site report in App. B.3]	Drs. Tyler B. Thompson, Dawn L. Shiang	http://www.dow.com/
Dow-Corning Corporation	Polysiloxanes, silicon-modified organics; optical & miscibility properties; permeation of gases; adhesion and interfacial interactions of siloxane networks; chemical reactivity and catalysis; environmental and cosmetics applications [See site report in App. B.3]	Drs. J. Cannady, C. Peter Qian,. Grigoras	http://www.dowcorning.com/
DuPont Pharmaceuticals Company	Pharmaceutical discovery and development [See site report in App. B.3]	Drs. S.L. Brenner, C.P. Decicco, D.J. Underwood, G.P. Brady, A.J. Tebben, Z.R. Wasserman, C.-H. Hwang, P. Meenan, J.P. Wohlbach	http://www.dupontpharma.com/

Table 10.5 (continued)
Recent and Ongoing Commercial Modeling Activities in the United States

Location	Type of activity	Contact(s)	Web links and references
E. I. du Pont de Nemours and Company, Central Research and Development	Homogeneous metallocene catalyst, hydrofluorocarbons, thermochemistry, crop-protection chemicals, support of combinatorial chemistry, identification of leads for agricultural chemicals, nylon, polymer photodecomposition [See site report in App. B.3]	Drs. Kerwin D. Dobbs, Paul D. Soper, Robert D. Waterland, John van Stekelenborg, Daniel A. Kleier, Ya-Jun Zheng, Laurie A. Christianson	http://www.dupont.com/
Eastman Chemical	Homogeneous and heterogeneous catalysis, thermochemistry, organic reaction mechanisms [See site report in App. B.3]	(Work discontinued, 1999)	http://www.eastman.com/
Eastman Kodak Company	Ground and excited electronic states of organometallic light-absorbing materials; T_gs by QSAR for a series of aryl amines (electroactive glasses)	Dr. John M. McKelvey and J.M. O'Reilly	http://www.kodak.com/; http://www.msi.com/materials/cases/qsprTg.html
Exxon (now ExxonMobil)	In 1980s, condensed matter physics; presently, lubricants	Dr. Ioannis Androulakis	http://www.exxonmobil.com/
Ford Motor Company	Materials and catalysis modeling, lubricants, atmospheric chemistry and hydrofluorocarbon coolants, UV-A photostabilizers, NOx SCR catalysis, alumina surface chemistry	Drs. Ellen B. Stechel, William F. Schneider, Michael L. Greenfield	http://www.ford.com/
Fujitsu, CAChe Software Group	Developers of CAChe semi-empirical software suite; formerly part of Oxford Molecular and of Tektronix [See Fujitsu site report in App. B.2]	Dr. David Gallagher	http://www.cachesoftware.com/
Gaussian, Inc.	Development and distribution of quantum chemistry codes	Dr. Michael J. Frisch	http://www.gaussian.com/
General Electric Corporation	Modeling of polycarbonate and polyester properties	Dr. John M. Whitney	http://www.ge.com/
General Mills, Inc.	Molecular modeling	Dr. Shen-Shu Sung	http://www.generalmills.com/
Halliburton Energy Services, Inc.	Design SandWedge™ additive to aid propping of hydrofractures in oil-field recovery	Dr. James Weaver	http://www.halliburton.com/hes/hes.asp; http://www.msi.com/materials/cases/sandwedge.html
Hercules Corporation	Specialty chemicals; e.g., modeling of polysaccharide rheology	Dr. Robert Grasso	http://www.herc.com/

Table 10.5 (continued)
Recent and Ongoing Commercial Modeling Activities in the United States

Location	Type of activity	Contact(s)	Web links and references
Hoechst Celanese (now Celanese)	Screening of miscible polymer pairs; polymers for gas separation membranes	Dr. S.H. Jacobson	
HRL Laboratories	Modeling III-V compound semiconductors	Dr. Mark Gyure	http://www.hrl.com/TECHLABS/isl/isl_home.shtml
IBM, Almaden Research Center, San Jose, CA	Triphenylamine as a hole transport molecule, decomposition of poly(perfluorinated ethers); Mulliken *ab initio* and molecular simulations code [See IBM-Zürich site report in App. B.1]	Dr. Jacob Pacansky	http://www.hirao.com/mulliken/prod/product.html; *Bull. Chem. Soc. Japan* 70, No. 1, 55-59 (1997); *Chemistry of Materials* 8, No. 12, 2788-2791 (1996)
IBM, T.J. Watson Research Center, Yorktown Heights, NY	Algorithms for parallel-computer molecular dynamics; Blue Gene special-purpose petaflop computer for protein folding [See IBM-Zürich site report in App. B.1]		http://www.research.ibm.com/news/detail/bluegene.html
Lubrizol Corporation	Specialty automotive chemicals, lubricant additives		http://www.lubrizol.com/
Lucent Technology, Bell Laboratories	Materials and reaction modeling for microelectronics (being spun off in 2001 into new company, Agere Systems), organic dyes, semiconductor lasers, sol-gel processes; develop methods (G-3) and potential functions [See site report in App. B.3]	Dr. Conor S. Rafferty	http://www.lucent.com/
Marathon Oil Company (part of USX Corporation)	Hydrogenation catalysis, liquid-liquid extraction (discontinued in late 2000) [See site report in App. B.3]		http://www.marathon.com/
Merck & Co., Inc. - Merck Research Laboratories	Pharmaceutical discovery and development [See site report in App. B.3]	Dr. Simon Kearsley	http://www.merck.com/
Millennium Pharmaceuticals, Inc.	Diagnostics and drug development by computational chemistry, genomics-based information analysis, and high-throughput screening		http://www.mlnm.com/
Mobil (now ExxonMobil)	Absorption of polycyclic aromatic hydrocarbons across the skin correlated by SAR; had molecular modeling group in early 1990s	Dr. T.A. Roy	http://chemweb.com/library/gbhap/display.exe?jcode=224&action=render&rendertype=abstract&uid=224.I224T981004&iid=2:8:3619
Motorola, Inc.	Materials and process modeling for electronics applications [See site report in App. B.3]	Drs. Wolfgang Windl and Anatoli Korkin	http://www.motorola.com/

Table 10.5 (continued)
Recent and Ongoing Commercial Modeling Activities in the United States

Location	Type of activity	Contact(s)	Web links and references
Network Science Corporation	Web-based provider of information about computational chemistry for biotechnology, chemical, and pharmaceutical industries	Dr. Allan Richon and Merry Ambos	http://www.netsci.org/
NutraSweet Kelco Co. (part of Monsanto Co.)	Rheology of polysaccharide thickeners	Dr. Todd Talashek	http://www.nutrasweetkelco.com/; http://www.msi.com/materials/cases/nskelco.html
Pharmacopeia, Inc. and its subsidiaries Molecular Simulations, Inc. (MSI), Oxford Molecular Group, Synopsys Scientific Systems	Experimental and modeling tools for pharmaceuticals, chemicals, and materials [see site reports for MSI (emphasizing chemicals and materials) and Pharmacopeia Labs (emphasizing drug discovery) in App. B.3]	Drs. John J. Baldwin, Mike Stapleton, Peter Gund, Kenneth M. Merz, Bruce E. Eichinger	http://www.pcop.com/; http://www.msi.com/; http://www.oxfordmolecular.com/; http://www.synopsys.co.uk/
Phillips Petroleum Company / Chevron Phillips Chemical Company	Metallocene catalysts for making polyolefins [See site report in App. B.3]	Dr. Donald E. Lauffer	http://www.phillips66.com/
Praxair	Modeling for separations of industrial gases		http://www.praxair.com/
Procter & Gamble Company	Modeling of surfactant behavior	Dr. John C. Shelley	http://www.pg.com/; *J. Physical Chemistry B* 102, 6318-6322 (1998)
Q-Chem, Inc.	Development and distribution of Q-Chem (*ab initio* code)		http://www.q-chem.com/
Rohm and Haas Company	Paintings and coatings, electronics, household products (detergents and personal hygiene), water treatment, agricultural chemicals, adhesives, and plastics [See site report in App. B.3]	Dr. Susan J. Fitzwater	http://www.rohmhaas.com/
Schrödinger, Inc.	Software including Jaguar quantum-chemistry and MacroModel MM / MD codes	Thomas A. Halgren; R.A. Friesner, B.J. Berne, W.A. Goddard III, W.L. Jorgensen	http://www.schrodinger.com/
Shell Development Co.	Thermochemistry, product and process development support, homogeneous catalysis	Drs. John Walsh and Nick Gonzales	http://www.shell.com/
SSCI, Inc.	Work with Bristol-Myers Squibb to identify crystal forms in a two-component mixture	A.W. Newman	http://www.msi.com/materials/cases/bms-rietveld.html; http://www.ssci-inc.com/
Solutia, Inc.	Polymer modeling	Dr. Genzo Tanaka	http://www.solutia.com/
Texas Instruments	Modeling oxygen vacancy defect in silicate gate oxides	Drs. R.M. Wallace and S. Tang	http://www.ti.com/

Table 10.5 (continued)
Recent and Ongoing Commercial Modeling Activities in the United States

Location	Type of activity	Contact(s)	Web links and references
Union Carbide, now part of Dow	Homogeneous catalysis	Dr. Jack Smith	http://www.unioncarbide.com/
UOP	Diffusion in zeolites, crystallographic analysis from modeling X-ray diffraction	Dr. Robert L. Bedard	http://www.uop.com/
Vertex Pharmaceuticals	Development of new small-molecule drugs using computer modeling and combinatorial chemistry; Agenerase™ HIV protease inhibitor		http://www.vpharm.com/; http://www.atom-by-atom.com/
Wavefunction, Inc.	Developers of SPARTAN and MacSpartan computational chemistry software	Dr. Warren Hehre	http://www.wavefunction.com/
W.R. Grace & Company	Modeling of transport in zeolites, screening ZSM and faujasite for different uses (past work)	Dr. Robert L. Harding	http://www.grace.com/

For several reasons, these tables are inevitably incomplete, but one intentional omission deserves comment. While the international tables (Tables 10.1 to 10.4) include government and academic laboratories, the U.S. table does not (Table 10.5). Work of U.S. government laboratories will be described later in this chapter and site reports on Sandia National Laboratories and the National Institute of Standards and Technology (Appendix B.3) provide representative detail. In academia, molecularly based modeling is dispersed among chemistry, physics, life sciences, materials science and engineering, and chemical engineering, involving far too many groups in the U.S. to be listed here.

It is interesting, however, to examine the chemical engineering (ChE) academic community. One reason is that chemical engineers represent a significant potential constituency for these methods and their results. Another is because interest and activity have grown up mainly in the last 20 years due to a mixture of intellectual and application opportunities.

Early modeling activities in this community grew out of ChE's activity in molecular thermodynamics, leading to leadership roles in molecular simulations for rheology and vapor-liquid equilibrium. In the early 1980s, Keith Gubbins, then of Cornell, organized the first ChE symposium in molecular simulations, but it drew little interest. By 1985 when the next ChE symposium was held, about twenty U.S. ChE academic research groups were involved in molecular simulations. At the same time, none was working in computational quantum chemistry. That number subsequently grew to fewer than five in 1990 and on to about thirty by the end of 2000. In this community, modeling has been used mostly for kinetics of homogeneous reactions and heterogeneous catalysis, A list of ChE academics involved in developing or using molecularly based modeling methods is available at http://www.comsef.aiche.org/, totaling about sixty U.S. ChE academic research groups at the start of 2001.

GOVERNMENT ACTIVITIES IN MOLECULARLY BASED MODELING

Governments have played indirect and direct roles in developing and applying these methods. Avenues include external funding of research by academics and industry, internal development at government laboratories, and projects jointly involving industry and government or promoting multi-industry collaborations. Although many university research projects are government-sponsored, they are not considered here. Some "institutes" are academic institutions, and even in these cases, they may be de facto government research laboratories. Without direct knowledge, the panel may not have correctly identified them as government laboratories.

European, Japanese, and U.S. activities are summarized below. The largest concentrations of activity are thus included, yet other governments do sponsor significant work. Several examples may be cited:

- The NRC Steacie Institute for Molecular Sciences of Canada's National Research Council is one such example. Its Theory and Computation Program is involved in research applications in chemistry, physics, and biology.
- The Instituto Mexicano del Petróleo (IMP) provides an example of a government-owned company, Mexican Petroleum Corporation (PEMEX), that is conducting research involving molecularly based modeling. A program led by Dr. Marcelo Lozada emphasizes R&D using molecular simulations.
- The Venezuelan Institute of Scientific Investigations (IVIC) is a government-sponsored national laboratory whose Chemistry Center has grown out of earlier programs in biomedical research, nuclear engineering, and petroleum-related research.
- National laboratories in India include the National Centre for Biological Sciences, Bangalore, and the National Chemical Laboratory, Pune, which carry out molecularly based modeling in protein folding and catalysis, respectively.

European Government Activities

This section focuses on Europe-wide activities; on French activities of C.N.R.S. and the Institut Français du Pétrole; and on British activities as seen from the perspective or involvement of Daresbury Laboratory. Of course, other European governments individually fund work in computational chemistry. For example, the two-university Center for Atomic-Scale Materials Physics (CAMP, see site report) was established with the aid of the Danish National Research Foundation. However, the panel was most knowledgeable about these cases. In addition, many European governments sponsor academic institutions, laboratories and institutes, as well as research there. The distinction could not always be clearly interpreted by the panel, as in the case of the large number of Russian institutes reported as working in this area.

The following analysis gives the most coverage to chemicals and materials programs, although it is not intended to neglect life sciences. It should be noted that many topics, especially large-scale molecular modeling and solvation, are quite relevant in all these fields.

Europe-wide Activities

The European Council, made up of the leaders of the European Union (EU), has stated the goal of Europe's becoming "the most competitive and dynamic knowledge-based economy in the world" (Lisbon Portugal, March 23-24, 2000). The Council recognized and committed the European Union to sponsoring research that would achieve this goal. Organizationally, much of this research is sponsored by the European Union/European Commission through a series of "Framework Programmes for Research and Technological Development." Approximately 250 "Training and Mobility of Researchers" (TMR) networks were set up as part of the Fourth Framework Programme (1994-1998). The TMR programs have had four main components:

- Research Training Networks (RTN), several of which will be discussed below
- Access to Large-Scale Facilities, such as the 1996-2000 Icarus Supercomputer project at CINECA in Italy (http://www.cineca.it/indexe.html), in which Prof. Nicholas C. Handy of the University of Cambridge and Prof. Erio Tosatti of S.I.S.S.A. helped coordinate computation in chemistry and physics
- "Marie Curie" Research Training Grants, for example the post-doctoral work of Dr. Theodora Spyriouni at Rhône-Poulenc/Rhodia (see Appendix B.1)
- Conferences, short courses, and summer schools, as at CECAM (discussed below and in Appendix B.1)

These networks were set up for three or four-year durations, beginning during 1996-1998 and ending during 2000-2002. There are typically five to 12 different research groups involved in each RTN. Approximately 300 Research Training Networks are supposed to be funded in the Fifth Framework Programme (1999-2002, http://www.cordis.lu/), variously starting in 2000-2002 and ending in 2003-2006. A central part of these efforts is to train doctoral and post-doctoral researchers.

Two recently concluded, chemistry-focused RTNs were on "Quantum Chemistry of the Excited State" and "Potential Energy Surfaces for Molecular Spectroscopy and Dynamics." Others are in progress on "Molecular Design of Functional Liquid Crystals" and "New Routes to Understanding Polymer Materials using Experiments and Realistic Modelling." Each network draws together interest-linked academic groups from around Europe, both for developing and applying theory, codes, and methods.

Much of the work in materials physics is loosely coordinated through Daresbury Laboratory as the Psi-k Network, joining activities sponsored by the European Union/European Commission and the European Science Foundation. Three EU TMR networks are involved: "*Ab initio* calculations of magnetic properties of surfaces, interfaces and multilayers," "Electronic structure calculations of materials properties and processes for industry and basic sciences," and "Computational magnetoelectronics." These projects are complemented by the ESF STRUC program, "Electronic structure calculations for elucidating the complex atomistic behaviour of solids and surfaces," and the British Collaborative Computational Project CCP-9, "Electronic Structure of Solids." Although apparently not affiliated with the Psi-k network, another EU materials RTN project is on "Development, Validation and Application of Hybrid Atomistic Modelling Schemes."

The European Science Foundation (http://www.esf.org/) represents 67 research-funding organizations from 24 countries of Europe. Its funding goes largely to fund conferences and fellowships. The STRUC program was noted above. Another program, the SIMU program "Challenges in Molecular Simulations: Bridging the time-scale and length-scale gap," emphasizes development of tools to perform multiscale molecular simulations. About 140 laboratories participate.

CECAM (Centre Européen de Calcul Atomique et Moléculaire; see site report in App. B.1) is one of the major assets of the European community in this field, providing stimulating workshops, short courses, and conferences for the molecular modeling community. It has been effective at generating new ideas and also at engaging newer researchers in using these models. Many European leaders are concerned about being behind the U.S. and Japan in electronic materials. However, Europe has a very strong position in modeling condensed matter, as shown in its lead position in modeling heterogeneous catalysis using electronic structure calculations. CECAM has proven extremely important at helping the one-to-one transfer of information and fostering the collaborations that have created this lead.

In summary, most Europe-wide funding aims to stimulate short-term progress by drawing together top researchers from different institutions. These teams also represent multiple countries within the EU and often are a mixture of theorists and experimentalists. The other purpose, educating student researchers, is not purely long-term. The most effective form of technology transfer is the individuals themselves. Today's students will be next year's agents of transfer, as well as the next decade's leaders.

French-sponsored Activities

The panel identified French government-sponsored modeling in the National Center for Scientific Research (Centre National de la Recherche Scientifique, or CNRS), the Institut Français du Pétrole (IFP), and the CEA (Commissariat à l'Énergie Atomique). A relevant CNRS/IFP project is the GdR 12090 project on modeling of catalysis.

The heart of government-sponsored research in France is the CNRS (http://www.cnrs.org/). Founded in late 1939 as a central organization for government-sponsored basic and applied research, it focused solely on basic research after World War II, until the 1970s. Government-sponsored applied research was shifted to ORSTOM (French Institute of Scientific Research for Development and Cooperation), CNET (National Telecommunications Center), and the CEA (Atomic Energy Commission), but in the 1970s, engineering sciences were added to the CNRS activities.

The other defining step for the CNRS was its 1966 creation of "joint research units." These are university-based, CNRS-funded labs that have both university and CNRS researchers. In this sense, their closest parallel in the United States is probably Lawrence Berkeley Lab (DOE), at which many U.C. Berkeley faculty members have positions.

Throughout France, CNRS employed over 25,000 people in 2000, about half of whom were researchers. Its 2001 budget is 16 million francs, approximately $2.2 billion. There are 1,200 CNRS service and research units around the country, intended to cover all fields of research.

As a relevant example of CNRS's approach, the panel visited the CNRS/Ecole Supérieure de Chimie Physique Electronique de Lyon joint research unit (see site report in Appendix B.1). Professors Jean-Marie Basset and Philippe Sautet, respectively, conduct experimental and modeling research on catalysis. The CNRS connection helps bring them in close contact with French and international industry, as well as providing important resources. Other CNRS molecular modeling includes work in laboratories at Châtillon (mesoscale modeling of crystal microstructure and plasticity) and Strasbourg (quantum chemistry, molecular simulations, and high-performance computing for catalysis and biochemistry).

The French Petroleum Institute (Institut Français du Pétrole or IFP) is a center for research and education for the oil, natural gas and automotive industries (see site report in Appendix B.1). Its financial support comes primarily from the government, but it is closely tied to the industries it supports and to international activities. In applications of molecular modeling, it has been a leader for France. Both fundamental and applied research is carried out relevant to the fields that the IFP supports.

One project in which IFP plays a leadership role is GdR 12090, "Dynamique Moleculaire Quantique Appliquee a la Catalyse" (Quantum Molecular Dynamics Applied to Catalysis; see site report in Appendix B.1). This is a research consortium sponsored by CNRS, but it includes the IFP, members from outside France (T.U. Vienna, Eindhoven Univ. of Technology), a multinational industry (Totalfina), and CNRS labs and universities (Villeneuve d'Ascq, Poitiers, Nancy, Université Pierre et Marie Curie, and ISMRA in Caen). When this project began in 1994, it expected to use the Car-Parrinello codes, but the VASP code has proven to be more useful. Consortium members have assigned duties, distributed among theory, modeling, and numerous applications.

There is also molecular modeling in laboratories of the CEA (Commissariat à l'Énergie Atomique). Work includes polymer modeling at CEA-Le Ripault, potential-function development at CEA-Saclay, and condensed-matter physics at CEA-Grenoble.

British-sponsored Activities

The United Kingdom sponsors research on molecularly based modeling and its applications at Daresbury Laboratory (see site report in App. B.2 for details) and academic laboratories. It also sponsors a series of successful Collaborative Computational Projects (CCP-1, CCP-2, etc.), and a strategic-planning activity, the Foresight program. As described above, it participates actively in the European Science Foundation and other Europe-wide networks.

The UK government's Foresight Programme was created to guide government investment in the UK science and engineering base R&D. Details can be found at http://www.foresight.gov.uk/. This program, which began in 1994, is the first time that UK government has sat down with academics and industry to try to structure/focus fundamental research (similar to the American "Vision 2020" program described below). There are about 15 industry sector reports available from the first cycle of Foresight, and a further 12 were published in December 2000. All of these are available on the Web site.

Japanese Government Activities

The 1999 organization of Japanese government funding is summarized in a report in Appendix B.2. Three components of the government dominate this area: the Ministry of Economy, Trade, and Industry (METI, formerly the Ministry of International Trade and Industry, or MITI), Science and Technology Agency (STA), and the Ministry of Education, Science, Sports and Culture (Monbusho)[3]. Structural changes in the specific organizations do not change the picture that the Japanese government is actively supporting this technology as having present and promising value.

[3] Editor's note: STA and Monbusho are in the process of merging in 2001.

METI's focus is to foster industrial strength, especially by aiding industry-initiated consortia and industry collaborations with academia. Many of these projects are supervised by its New Energy and Industrial Technology Development Organization (NEDO). For example, NEDO supplies 90% of the funding for the Angstrom Technology Partnership (ATP) involved with JRCAT, discussed below.

One such NEDO project is "Platform for Designing High Functional Materials" at Nagoya University, widely referred to as the "Doi Project" after its leader Prof. Masao Doi (see site report in Appendix B.2). Technically, the Doi Project's goal is to develop computational methods for multi-scale modeling of polymer structure and properties, encompassing theory, algorithm, and code development. The project is formally carried out by the Japan Chemical Innovation Institute (JCII) and has involved 12 companies. Each company places one of its employees into the Doi team for the four-year span of the project (1998-2002), their salaries primarily funded by METI with some support from the Ministry of Education. Nagoya and six other university teams receive 5% of the funds. While companies are plainly interested in the code itself for their own projects, the government recognizes that the interacting team members will become the polymer-modeling infrastructure of Japan's future.

A second METI agency is the Information Technology Promotion Agency (IPA), whose Code Development Projects are important to Japan's activities in this field. Perhaps the most relevant projects are the CAMP and AMOSS projects. These are described in the NEC site reports because NEC staff chair the projects, although many companies are involved. Several individual researchers expressed their beliefs that code development was being sponsored partly out of concern that Microsoft and other non-Japanese software companies might otherwise come to dominate the field completely and subsequently serve Japanese users poorly.

Also part of METI is AIST, the National Institute of Advanced Industrial Science and Technology (formerly the Agency of Industrial Science and Technology). It has 15 institutes, including the National Institute for Advanced Interdisciplinary Research (NAIR) in Tsukuba, described in the JRCAT site report. NAIR and the "Angstrom Technology Partnership" jointly set up the Joint Research Center for Atom Technology (JRCAT; see site report). It was formed to help carry out the 1992-2002 "Ultimate Research and Development of Manipulation of Atoms and Molecules (Atom Technology Project)."

The Japanese Science and Technology Agency sponsors mostly academic research grants through its "Fund of Promotion of Science," "Strategic Basic Research Promotion Program," and "Research and Development Applying Advanced Computational Science and Technology" program. Some national laboratory and industrial research is also supported. An example is RIKEN, the Institute of Physical and Chemical Research, which conducts fundamentals-focused research in many scientific areas (see site report). It is a semi-public corporation, completely supported by the Science and Technology Agency.

A third large agency is the Ministry of Education, Science, Sports and Culture (Monbusho), especially through its "Research for the Future" program. One of its grants is described in the Institute for Molecular Sciences site report in Appendix B.2.

The Okazaki National Research Institutes include the Institute for Molecular Science (see site report), one of the world's top programs in molecularly based modeling. It operates a supercomputer center that provides much of the computer time used by academic chemistry- and biology-based modeling in Japan.

In Japanese research on condensed-matter physics, the Institute for Solid State Physics (ISSP) of the University of Tokyo is a central resource (see site report). The ISSP was established on April 1, 1957, upon the recommendation of the Science Council of Japan and with the concurrence of the Ministry of Education and the Science and Technology Agency. Its supercomputer center provides computational power for many projects in condensed-matter modeling for universities throughout Japan—200 research groups located at more than 60 Japanese universities.

U.S. Government Activities

There is significant activity in molecularly based modeling by U.S. government departments and agencies, including those who requested the present study. The National Science Foundation (NSF) contracted RAND

to search the FY1998 government-funding RaDiUS database for sponsored research projects in this field. The method and results of analysis are detailed in Appendix D. Estimating the molecular modeling content based on keywords, in FY1998 the U.S. government was estimated to have funded $160 million in grants and contracts related to molecular modeling research. NSF provided a more detailed breakdown, discussed below.

Activities are described below more specifically, organized by agency.

Department of Commerce

The National Institute of Standards and Technology (NIST, Gaithersburg, MD) is the center of research and funding for the Department of Commerce; more details are provided in a site report in Appendix B.3. Molecularly based modeling is incorporated into a number of its research programs. Leaders are as follows:

- Center for Advanced Research in Biotechnology (Chemical Science and Technology Laboratory)
- Computational Chemistry Group (Chemical Science and Technology Laboratory)
- Experimental Kinetics and Thermodynamics Group (Chemical Science and Technology Laboratory)
- Center for Theoretical and Computational Materials Science (Materials Science and Engineering Laboratory)
- Fire Sciences Division (Building and Fire Research Laboratory)
- Information Technology Laboratory

NIST's Advanced Technology Program has provided matching funds with industry for a number of projects that are high-risk or that build infrastructure. Examples include development of DFT software for biological molecules (Biosym Technologies, Inc., 1992-95); analysis of methaluminoxane chemistry (MAO) and development of new metallocene activator (Albemarle Corp., 2000-2003); and mechanistic modeling of metallocene catalysts for polyolefins (Phillips Petroleum, 1995-98; see Phillips site report in App. B.3). Proposed new activities related to this area would include "Combinatorial Chemistry and Materials Research" and "Initiatives for Bioinformatics." The strategy of the ATP program has been politically controversial, so its long-term future is uncertain.

Department of Defense (DOD)

An important activity that involves all the branches of the Department of Defense is the High-Performance Computing Modernization Program (http://www.hpcmo.hpc.mil/). There are three parts to the programs: high-performance computing centers, the Defense Research and Engineering Network (DREN), and the Common High Performance Computing Software Support Initiative (CHSSI). The Army's PET program (Programming Environment and Training; http://www.arl.hpc.mil/PET/) is a related component. CHSSI is broken into ten computational technology areas, including computational chemistry and materials science (http://www.hpcmo.hpc.mil/Htdocs/CTAs/ccm.html). Within CCM, there are several projects:

- Car-Parrinello code development: DOD Planewave and ACRES codes, development led by Dr. David J. Singh of the Naval Research Laboratory. Both codes are publicly available and free over the Web. DOD Planewave has also been adapted for Windows computers by Alex Voit of the Institute of Chemistry FEBRAS in Vladivostok, Russia, and this code is also being distributed free.
- Parallel quantum-chemistry code development using GAMESS-US.
- Tight-binding molecular dynamics code, developed by Prof. Furrukh S. Khan and Dr. Florian Kirchhoff at Ohio State University. TBMD and related codes are distributed free.
- Classical molecular dynamics: development of the parallel FMD code, which incorporates the three-dimensional fast multipole method for order-N calculational speed.

An annual Challenge Proposal program distributes large blocks of supercomputing time for DOD contractors, labs, and grantholders. These represent Grand-Challenge-level applications of high-performance computing.

Defense Advanced Research Projects Agency (DARPA). Several projects are sponsored by DARPA to move existing and new codes to parallel computers. One such project is "Portable Parallel Conditions," carried out by Xerox Palo Alto Research Center, Oak Ridge National Laboratory, and the University of Tennessee (http://sandbox.xerox.com/gilbert/ppp/). Its goal is to develop sparse matrix solvers that can be used easily to speed the writing of parallel codes, including for computational chemistry. In another project, IBM worked with a number of independent software vendors, including MSI, to parallelize their engineering-related codes. MSI was able to commercialize MPI parallel versions of Discover and its Dmol DFT code. The latter was tested as reaching 97% parallel efficiency on 16 processors.

Air Force Office of Scientific Research (AFOSR) and Aeronautics Systems Center (ASC). The Molecular Dynamics and Theoretical Chemistry Programs (Dr. Michael R. Berman, program manager, http://afosr.sciencewise.com/afr/afo/any/menu/any/BAA2000.HTM#3c) sponsors work on basic understanding of reactions and energy transfer. By fostering interactions among principal investigators, Air Force laboratory scientists, and industry representatives, this basic understanding is applied to technological problems: propulsion, energetic materials, diagnostics, atmospheric chemistry, lasers, ceramics, and metals.

At Wright-Patterson Air Force Base, the Aeronautics Systems Center (http://www.asc.hpc.mil/) is Air Force's Major Shared Resource Center involved in computational chemistry and materials science.

Army Research Office (ARO) and Army Research Laboratory (ARL). The ARO sponsors work in basic and applied physics and chemistry, including molecularly based modeling in physics and chemistry. Areas of interest in physics include modeling of alloy strength, electronic band structure, reaction kinetics, nonequilibrium dynamics, and quantum optics. Its Chemical Sciences Division (Dr. Robert Shaw) funds theoretical chemistry directly, as well as work in reaction kinetics, polymers, electrochemistry, catalysis, and organic chemistry for protection from toxic materials that may use computational chemistry.

A major activity of ARL is its partnership with the Ohio Supercomputer Center (http://www.osc.edu/) as a center for its Computational Chemistry and Materials Science (CCM) activities in the PET program (Programming Environment and Training, http://www.arl.hpc.mil/PET/).

Office of Naval Research and the Naval Research Laboratory. ONR is the parent organization of the NRL, but it also funds a large amount of applied research in chemical and physical sciences and engineering (http://www.onr.navy.mil/sci_tech/engineering/).

The NRL's Center for Computational Materials Science (http://cst-www.nrl.navy.mil/), led by Dr. Dimitrios A. Papaconstantopoulos, conducts theoretical and computational research on modeling of materials, tied closely to experimental data. The Theoretical Chemistry Section in the Chemistry Division, headed by Dr. Carter T. White, develops theory and codes for diverse applications ranging from combustion to surface science ((http://www.chemistry.nrl.navy.mil/6180/6189.html)).

Department of Energy (DOE)

The DOE has been contributed to many developments in molecularly based modeling through project research funding and its national laboratories.

Recently, the Office of Industrial Technologies (http://www.oit.doe.gov/) has sought to aid energy efficiency by partnering with the nine most energy-intensive industries: agriculture, forest products, mining, aluminum, glass, petroleum, chemicals, metal casting, and steel. Molecular modeling has cross-cutting value through all these industries. To develop long-term plans for action, OIT embarked on a series of workshops and "technology roadmaps" under the title "Vision 2020." Several reports for the chemical industries (http://www.oit.doe.gov/chemicals/page9.shtml) touch on uses of computational chemistry, including materials technology, catalysis and reaction engineering, new process chemistry, and a study specifically focused on the subject (http://www.oit.doe.gov/chemicals/pdfs/chemvision2020.pdf). To implement this vision, OIT has also supported R&D grants with substantial industry participation and co-funding.

The Office of Science (http://www.er.doe.gov/) has aided work in theory, applications, and communication of the results. Basic Energy Sciences supports contracts for basic research in chemical physics and condensed-matter physics to principal investigators both within and outside the national-laboratory structure.

At Lawrence Livermore, Los Alamos, and Sandia National Laboratories, DOE's Accelerated Strategic Computing Initiative (http://www.sandia.gov/ASCI/) has developed the fastest supercomputers in the world (http://www.llnl.gov/asci/platforms/). ASCI was created to develop computational simulation that would substitute for testing of nuclear weapons. That has led to the creation of ASCI White (12.3 teraflops, IBM RS/6000, LLNL), ASCI Blue Pacific (LLNL, 3.9 teraflops) and Blue Mountain (LANL, 3.9 teraflops), ASCI Red (2.4 teraflops, SNL), and ASCI Cplant (1 teraflop, projected to rise to reach 1 petaflop by 2010). By 2004, the government intends to build a 100-teraflop computer. Along with hardware, ASCI is funding software development, including software for computational chemistry (see Sandia site report in App. B.3).

Ames Laboratory. The Condensed Matter Physics Group (http://cmp.ameslab.gov/) conducts research, develops codes, and helps lead several DOE initiatives in computational materials science. Its research focus is modeling large molecular systems using various tools including its AL_CMD code (Ames Lab Classical Molecular Dynamics). This code is designed for portable use on massively parallel computers and is freely distributed.

Two of its wider roles are of general interest. It hosts the Computational Materials Sciences Network (http://cmpweb.ameslab.gov/ccms/) sponsored by DOE's Division of Materials Sciences, Condensed Matter Physics & Materials Chemistry Programs. This network draws together government, academic, and industrial researchers seeking to solve materials problems by bridging disciplinary and organizational boundaries. Second, Ames is part of a "Partnership In Computational Science" (PICS) consortium on "First-Principles Simulation of Materials Properties" (http://www.ccs.ornl.gov/GC/materials/MShome.html) begun in 1994 involving Brookhaven National Laboratory, Oak Ridge National Laboratory, and six academic institutions. The consortium has focused on building and parallelizing a hierarchy of models using classical potentials, tight-binding models, and *ab initio* methods.

Argonne National Laboratory (ANL). Argonne has long been a leader in modeling of chemistry and materials. Dr. Aneesur Rahman of its Solid State Division (now the Materials Science Division, http://www.msd.anl.gov/) is generally credited with the development of modern molecular dynamics, simulating a system of 864 argon atoms on a CDC 3600 computer in 1964. Its Chemistry Division conducts coordinated gas-phase experimental and theoretical analyses of state-selective chemistry, reaction kinetics, and thermochemistry of photoionization spectroscopy; calculation of high-level potential energy surfaces; and development of COLUMBUS (http://chemistry.anl.gov/chem-dyn/chemdynreview/col.html), a system of codes for electronic structure calculations. A Molecular Materials/Computational Materials Chemistry group, operated by both divisions and led by Dr. Larry Curtiss, carries out research including development of the G2 and G3 test sets and methods (http://chemistry.anl.gov/compmat/).

Brookhaven National Laboratory (BNL). The Brookhaven Center for Data Intensive Computing includes projects on quantum mechanical modeling of molecular spectroscopy and of magnetic and supercomputing oxides (http://www.bnl.gov/cdic/). The Protein Data Bank of structures that was long maintained by Brookhaven was shifted in 1999 to the Research Collaboratory for Structural Bioinformatics (RCSB; http://www.rcsb.org/), a consortium composed of Rutgers University, the University of California at San Diego, and the National Institute of Standards and Technology. RCSB was funded by a $10 million, five-year award from NSF, DOE, and two units of the National Institutes of Health—the National Institute of General Medical Sciences (NIGMS) and the National Library of Medicine (NLM).

Lawrence Berkeley National Laboratory (LBNL or LBL). The Scientific Computing Group in LBL's National Energy Research Scientific Computing Center (http://www.nersc.gov/) conducts research on computational chemistry and physics. Molecularly based modeling is carried out in the Chemistry Division and Materials Science Division.

LBL has been the site of the National Resource for Computation in Chemistry (NRCC) since 1978. There, in 1980, Dr. Michel Dupuis and co-workers developed the *ab initio* electronic-structure code General Atomic and Molecular Electronic Structure System (GAMESS). This continues to be a widely used code, now

available in separate versions in the United States (http://www.msg.ameslab.gov/GAMESS/GAMESS.html) and the UK, as GAMESS-UK (http://www.dl.ac.uk/TCSC/QuantumChem/Codes/GAMESS-UK/) distributed by Computing for Science, Ltd. (http://www.dl.ac.uk/CFS/). Dr. Dupuis is now part of the theory and code development effort at Pacific Northwest National Laboratory. Prof. William A. Lester, Jr., developer of the Quantum Monte Carlo electronic structure method (http://www.cchem.berkeley.edu/~walgrp/) and the code Quantum MagiC, was director of the NRCC from 1978-81 and continues theoretical chemistry work at LBNL and the University of California, Berkeley.

Lawrence Livermore National Laboratory. LLNL carries out molecularly based modeling in chemistry, condensed-matter physics, and biology. It is the home of the ASCI White and Blue Pacific supercomputers.

Work in the Chemistry and Chemical Engineering Division of LLNL's Chemistry and Materials Science Directorate (http://www-cms.llnl.gov/) is focused on applications in three areas:

- Energetic materials (detonation and slow-burning conditions, kinetics of high-energy density materials, and electronic structure of high explosives)
- Combustion chemistry (developing mechanisms of elementary reactions for ignition, flame propagation, quenching and emissions from internal combustion engines)
- Destruction of chemical warfare agents

In the same directorate's Materials Science and Technology Division, the Computational Materials Science program is treated as essential both as part of the division's own basic and applied research, and also for resolving materials questions and needs elsewhere in LLNL. Techniques and applications include the following:

- Quantum mechanics (total energies, magnetic, electronic, thermodynamic and transport properties)
- Atomistic simulation applied to defects and diffusion in solids (radiation damage, ion implantation, dopant diffusion)
- Phenomenological modeling of processes (metal-working operations such as casting and welding; material failure such as crack propagation and fatigue)

Molecular modeling in the life sciences is concentrated in the Computational Biology Group (http://gutenberg.llnl.gov/~colvin/) of the Biology and Biotechnology Research Program (http://www-bio.llnl.gov/). Its subgroups are Computational Biochemistry (DNA operation, damage mechanisms, and repair—on enzyme reaction mechanisms) and the Protein Structure Prediction Center (methods of identifying protein structure from sequence; objective testing of methods via blind prediction).

Los Alamos National Laboratory. LANL is the home of the ASCI Blue Mountain supercomputer. Its Theoretical Chemistry & Molecular Physics Group (http://www.t12.lanl.gov/) carries out theory, method, and application development on materials modeling in the broad sense of gas, liquid, and solid modeling. Topics include electronic properties of molecules, dynamics and kinetics of chemical reactions, atomistic simulations of materials, molecular modeling of catalysts, the study of solute-solvent interactions, and chemical process modeling. The group also developed the LANL2DZ electronic basis set, which uses effective core potentials to capture the electronic structure of metals and other heavy atoms, including relativistic effects.

These tools are also being used in computational biology within the Bioscience Division (http://bdiv.lanl.gov) and in the Materials Science and Technology Division (http://www.lanl.gov/orgs/mst/).

Oak Ridge National Laboratory. ORNL's Chemical Technology Division has several programs focused on molecular-based science and technology (http://www.ornl.gov/divisions/ctd/computational.htm), including an emphasis on molecular simulations and parallel processing for separations, material processing, and supercritical-fluid waste incineration. This work includes a project on "Molecular-Based Study of Fluids and Amorphous Materials Using Parallel Supercomputers" headed by Prof. Peter T. Cummings (a member of this panel) who holds a joint appointment as Distinguished Professor of Chemical Engineering, Chemistry and Computer Science at the University of Tennessee and also as UT/ORNL Distinguished Scientist in the

Chemical Technology Division. His position represents another approach by which government sponsors and promotes work in this field.

Computational modeling of hard condensed matter is found in the Computational Modeling and Simulation Section of the Computational Physics and Engineering Division (http://www.cad.ornl.gov/).

A range of molecular modeling and informatics is carried out in the Computational Biology Section (http://compbio.ornl.gov/) of the Life Sciences Division. Work includes genetics research and system development in genomic sequencing, computational genome analysis, computational protein structure analysis, high-performance biocomputing, bioinformation systems, biosystems modeling, and genetics information management. They also provide user services for the general community, including computer-annotated genomes in Genome Channel and Genome Catalog.

Pacific Northwest National Laboratory. PNNL operates the William R. Wiley Environmental Molecular Sciences Laboratory (EMSL), which has a significant component in computational chemistry and high-performance computing. Two key parts are the Molecular Science Computing Facility (http://www.emsl.pnl.gov:2080/capabs/mscf/) and the Theory, Modeling, and Simulation Division (Dr. David A. Dixon, Associate Director; http://www.emsl.pnl.gov:2080/homes/tms/). The MSCF is a user facility as well as an internal resource. TM&S combines theoretical and computational chemistry, materials science and condensed-matter physics, computational molecular biology, computer science, and applied mathematics to address general environmental problems and those associated with clean-up of nuclear waste.

A Molecular Science Software Suite (http://www.emsl.pnl.gov:2080/capabs/mscf/software/ms3-1999.html, MS^3) has been developed and is evolving. It is comprised of Ecce (Extensible Computational Chemistry Environment, a graphical user interface and problem-solving environment), Parsoft (library of parallel computing software), and NWChem (electronic structure code for parallel computation). This software suite is being distributed free of charge.

Sandia National Laboratories. A detailed site report on Sandia is included in Appendix B.3. It is heavily involved both in developing computational tools and in applications from materials to genomics. Sandia is home to the ASCI Red supercomputer, the first computer to achieve teraflop performance, and its Cplant project of commodity-based-CPU parallel computing holds the goal of petaflop computing within this decade.

Department of Health and Human Services (HHS): National Institutes of Health

HHS conducts and supports a large amount of molecularly based modeling through the National Institutes of Health (NIH). Its Center for Molecular Modeling (CMM, http://cmm.info.nih.gov/) provides software, advisory, and collaborative services within NIH for modeling problems that arise in life sciences. In addition, it provides a number of Web-based resources to the larger user community, such as a comprehensive molecular modeling software list (http://cmm.info.nih.gov/software.html), databases, and literature lists.

Another significant NIH facility is the National Cancer Institute's supercomputing facility at Frederick, MD, the Advanced Biomedical Computing Center (ABCC, http://www-fbsc.ncifcrf.gov/). The facility provides state-of-the-art computing support and technology to NCI and NIH, as well as to extramural biomedical researchers.

Environmental Protection Agency

The EPA maintains its National Environmental Supercomputing Center (http://www.epa.gov/nescweb0/, http://www.epa.gov/nesc/15_services/comp_chem_info.html) at Research Triangle Park, NC. The facility is dedicated to environmental research by EPA scientists, managers, and qualified EPA customers, and computational chemistry services are offered through NESC's Scientific Visualization Center. Two full-time molecular modelers are on staff. The main codes used are Gaussian, SYBYL, MIDAS Plus, and AVS Chemistry Viewer. Contacts are Dr. Charles Foley and Dr. James Rabinowitz.

Earlier work includes work toward an electronic-structure database of industrial chemicals, computed by Dr. Steven Bradbury and co-workers (http://www.epa.gov/nesc/SCS/chem/) of the EPA Mid-Continent Ecology Division, Duluth, MN. Aquatic toxicity values and toxicodynamic responses were measured and related by QSAR methods to electronic-structure parameters calculated for 600 compounds (*Toxicol. Lett.* 79, 229-237, 1995). Other QSAR relations were developed for receptor binding and toxic response of polychlorinated biphenyls, polychlorinated dibenzofurans, and dioxins using MOPAC PM3 calculations (*Environmental Health Perspectives*, 104:12, December 1996; http://ehpnet1.niehs.nih.gov/docs/1996/104-12/mekenyan.html).

Work on computational chemistry modeling for massively parallel computers is also sponsored externally by the EPA through its High-Performance Computing and Communications program. Gao and colleagues at the State University of New York at Buffalo (now of the University of Minnesota) have been developing a quantum mechanics-molecular mechanics code, and Yang and colleagues at Duke University are developing a divide-and-conquer DFT code for modeling very large systems (http://www.epa.gov/HPCC/molmod.html).

National Aeronautics and Space Administration (NASA)

NASA conducts modeling of materials and biological molecules, notably at its Ames Research Center (http://www.arc.nasa.gov/) in California. Its Center for Computational Astrobiology (NCCA, http://cca.arc.nasa.gov/) is combining computational chemistry and biology with geology and astronomy to study the origins of life on Earth and its possibility on other planets. The range of projects is illustrated by interests of different researchers:

- Dr. Charles Bauschlicher, focusing on formation of polycyclic aromatic hydrocarbons in interstellar space
- Dr. Richard Jaffe, conducting a range of molecular-dynamics and quantum-chemistry calculations
- Dr. Timothy Lee, developing quantum-chemistry methodologies and applying theoretical spectroscopy
- Dr. Andreas Parusel, *ab initio* calculations for biological nanotechnology
- Dr. Deepak Srivastava, simulating properties of nanoscale materials
- Dr. Stephen Walch, computational chemistry applied to astrobiology, transition-metal chemistry, combustion chemistry, atmospheric chemistry, and chemical vapor deposition (CVD) reactors

National Science Foundation (NSF)

The NSF sponsors specific work on molecularly based modeling through several of its divisions. A search of NSF grants and contracts by keyword, augmented by NSF staff analysis, yields an estimate for FY98 funding in the field as shown in Table 10.6 (adapted from Appendix C).

Table 10.6
Estimates of NSF Funding Involving Molecularly Based Modeling, FY98 ($ million)

NSF Directorate	Estimate
Biological Sciences	$ 4,000,000
Computer & Information Science & Engineering	1,400,000
Education & Human Resources	175,000
Engineering	3,700,000
Geosciences	800,000
Mathematical & Physical Sciences	20,000,000
Social, Behavioral & Economic Sciences	150,000
TOTAL	**$30,000,000**

The Directorate for Mathematical and Physical Sciences (http://www.nsf.gov/home/mps/) provides the largest amount of NSF funding, especially through the Division of Chemistry

(http://www.nsf.gov/mps/chem/) and the Division of Materials Research (http://www.nsf.gov/mps/dmr/). Nanoscale science and nanotechnology is another field where these methods, extending from atomic clusters into the mesoscale, are naturally important. The National Nanotechnology Initiative (http://www.nano.gov) is headed by Dr. M.C. Roco of the NSF Engineering Directorate, where participating organizations also include other parts of NSF, the Departments of Agriculture, Commerce, Defense, Energy, Justice, and Transportation, and EPA, NASA, and NIH.

Direct funding for theory and modeling is easier to identify, but the use of molecularly based modeling extends into many scientific programs of NSF. In chemistry, much of the leading theoretical work in the United States is funded through the Theoretical and Computational Chemistry Program, while in materials research, the Materials Theory Program (http://www.nsf.gov/mps/dmr/theory.htm) plays a similar role.

Sponsorship of major computing centers is another role that NSF plays in this field. NSF's Supercomputer Centers Program operated five centers beginning in the mid-1980s. In 1997, this program was replaced with the NSF Partnerships for Advanced Computational Infrastructure program (http://www.cise.nsf.gov/acir/). PACI was formed with two groups ("alliances") of academic partners, the National Computational Science Alliance (NCSA) based at the University of Illinois and the National Partnership for Advanced Computational Infrastructure (NPACI, http://www.npaci.edu/), centered at the San Diego Supercomputer Center. The intent of each alliance is to use local and distributed supercomputing hardware to develop a national computational problem-solving environment.

NON-GOVERNMENT, MULTI-LATERAL ACTIVITIES

Professional societies have provided an important venue for communicating new developments and new applications among academics, government researchers, and companies. In the U.S., leading roles have been played by the American Chemical Society (ACS, http://www.acs.org/), American Physical Society (APS, http://www.aps.org/), and the American Institute of Chemical Engineers (AIChE, http://www.aiche.org/). The ACS divisions of Computers in Chemistry, Industrial and Engineering Chemistry, and Physical Chemistry have organized numerous symposia on theory, on computational chemistry in biological systems, and on industrial application. AIChE has maintained annual programming since 1985, building to topical conferences in 1998 and 2001 on "Applying Molecular Simulations and Computational Chemistry" and a Computational Molecular Sciences and Engineering Forum (http://www.comsef.aiche.org/) formed in 2000. Other meetings and workshops have been organized by the U.S. government, Gordon Conferences, and other organizations.

The Royal Society of Chemistry has played a similar role in the UK. Its Molecular Modeling Group (see site report in App. B.1) has provided leadership and a contact point for activities bridging industry, government and academia.

The international Molecular Graphics and Modelling Society (http://www.mgms.org/) has contributed to the development and application of molecular modeling and computational chemistry since 1981. It reports having about 500 members worldwide. Its activities include sponsoring conferences and the *Journal of Molecular Graphics and Modelling* (Elsevier, http://www.elsevier.nl/locate/jmolgraph) which focuses on computer-aided molecular design.

The Internet adds new means of displaying and sharing information, reflected in good measure by the Web links cited throughout this study. An important example for this field is the Computational Chemistry List (CCL, http://www.ccl.net/chemistry/), a free electronic discussion forum about technical and nontechnical aspects of computational chemistry. Created by Jan Labanowski, in January 1991, the CCL has been hosted, supported, and encouraged by the Ohio Supercomputer Center (OSC, http://www.osc.edu/). Despite (or with the aid of) a 1996-99 NSF grant of support, by 1999, the CCL consumed more than 90% of the Web and ftp traffic of the Ohio Supercomputer Center. Using figures at the beginning of summer 2000, 40% of CCL's subscribers were in the U.S. with the remaining 60% in over 50 other countries on all continents. CCL's subscribers were reported to come from educational institutions (50%), commercial (36%), government (5.5%), military (2%), and non-profit (1.5%), with the remainder being Internet service providers.

SUMMARY OF INTERNATIONAL COMPARISON

Table 10.7 represents a distillation of this panel's opinions, which are inevitably imperfect. While we recognize that there are significant activities in other parts of the world, the table restricts itself to consideration of activities in the U.S., Europe, and Japan.

Even in an era of instant global communications, distinct regional differences and strengths exist. Some differences are rooted in societal or governmental differences. However, most regional strengths are consequences of geographical centers of intellectual achievement.

The origins of these centers are individuals and small teams of gifted researchers. The strengths have evolved because government, industry, or individual commitment to specific ideas and approaches has fostered intellectual centers. They are regional, at least initially, because technology is transferred best by the movement of knowledgeable people.

At the same time, each region has benefited by people spending careers or extended periods across international borders. Japanese organizations have been among the most conscientious and successful in doing so. European countries have been especially effective at creating multi-national research and educational programs such as CECAM.

The United States leads this field in many scientific areas. However, Canada has particular strengths in DFT methods and homogeneous catalysis; Europe in heterogeneous catalysis, mesoscale, and materials modeling; and Japan in materials modeling and special-purpose computers. Major government-industry initiatives are underway in Europe and Japan, notably in multi-scale materials modeling and in development of chemistry-capable *ab initio* molecular dynamics codes. In European and U.S. assessments of nanotechnology, it was also concluded that to advance the field most quickly—and competitively—the need is acute for applying new and existing methods of molecularly based modeling.

The most pronounced technical differences proved to be between different industries. On technical matters, pharmaceutical companies in Europe, Japan, and the United States have much in common. The same can be said for the polymers industry, the specialty chemicals industry, the oil and gas industry, the consumer products industry or the microelectronics industry or the ceramics industry. Noting only a few example companies for each category, industries applying molecularly based modeling include the following:

- Pharmaceuticals (Merck, Novartis, Takeda Chemicals)
- Crop-protection chemicals (DuPont, Sumitomo Chemical, Bayer)
- Polymers, glass, and structural materials (Asahi Chemicals, Owens-Corning, Rhône Poulenc/Rhodia, W.R. Grace)
- Electronic and photonic materials (Motorola, Toshiba, Lucent)
- Homogeneous and heterogeneous catalysts (Ford, Haldor Topsøe, Ube Industries)
- Industrial gases (Advantica Technologies—formerly BG Technology, Air Liquide, Air Products and Chemicals)
- Personal-care, food, and consumer products (Colgate Palmolive, Unilever, Kellogg, 3M)
- High-volume chemicals and materials (Dow, BASF, Rohm & Haas)
- Dyes and pigments (Bayer, Mitsubishi)
- Films and imaging (Fuji Photo Film, Xerox)
- Fuels and automotive chemicals (Chevron, Totalfina, Lubrizol)
- Commercial software for calculations (MSI/Pharmacopeia, Gaussian, Fujitsu)

Table 10.7
Comparative status of Molecular and Materials Modeling in Europe, Japan, and the United States

Technology	Europe Status	Japan Status	US Status
Methods			
- Quantum Mech. & Hybrid Methods, Molecular	****	***	*****
- Quantum Mech. & Hybrid Methods, Condensed Matter	*****	****	****
- Molecular Simulations			
Biological systems	****	***	*****
Fluids	*****	***	*****
Solids, devices	***	****	***** (devices)
Polymers	*****	*****	****
Kinetics – Monte Carlo	***	**	***
Mesoscale	*****	***	***
- Solvation	*****	****	*****
- Informatics & Optimization	****	***	*****
- Hardware, Operating Systems, Math Methods	****	*****	*****
Applications			
- Chemical Process Development	****	****	****
- Chemical Process Design and Manufacturing	**	**	**
- Chemical Product Design Fuels, specialty chemicals, coatings, surfactants	****	****	****
- Reaction Thermochemistry & Kinetics			
Noncatalytic	****	***	*****
Homogeneous catalysis	***	*****	*****
Heterogeneous catalysis (mostly academic)	**** (QC)	***	**** (sim)
- Physical and Electronic Materials			
Metals, semiconductors	***	*****	*****
Ceramics, glasses	****	*****	****
Polymers	*****	****⇑	*****
- Biomolecules and Biologically Active Materials			
Pharmaceuticals	****	***	*****
Agricultural products	***⇑	?	***⇑
- Application to Analytical Measurements			
Chemical analysis	***⇑	?	****⇑
Material analysis	***⇑	**⇑	**⇑
Biological analysis	****	***	*****
Transfer of Science to Technology	****	*****	***

Key: Number of stars indicates approximate level and quality of activity by region;
0 => no activity, no quality; 5 => very large amount of activity or high level of quality;
up arrows (⇑) indicate increasing level of activity

To a surprising extent, infrastructure and organizational issues are the same in Europe, Japan, and the United States. The process and organization of these activities have proven as important as the science and technical implementation. Quantum mechanics and statistical mechanics are well-established fields. However, their technological applications have been held back by information technology (computers and software) and by how they are applied, including whether expectations are reasonable. Success and failure always depend on the right people working on the right problem with the right tools.

Raw computing power can't buy success, but advances in computing power, visualization, and information technology have made many new applications of molecular modeling feasible. Growth in parallel computation ability has not kept pace with advances in parallel computer hardware, largely because methods lag for construction and optimization of user codes. Ironically, using specialized parallel hardware (SGI Origins, for example) appears to be slowing as the economics of commodity CPUs drive rapid increases in their use (Beowulf clusters). User interfaces with software remain a problem, especially for generation of thermochemical and kinetics parameters.

Molecular modeling is now well established in many companies, while other companies are just beginning; and in others, it has proven vulnerable to reduction or elimination. Company by company, the question comes down to this: "Is molecular modeling considered a necessity or a luxury?" Management and employees must agree.

In pharmaceutical and chemical development, availability of molecular modeling software and/or specialists now is often considered to be as routinely necessary as having analytical instruments or analytical chemists. New chemists take for granted the usefulness of 3-D visualization or property estimations. So do experienced chemists who have found it beneficial in their work or their colleagues' work. People will seek out tools and ways of thinking that help them. Management also responds to tangible benefits, and possible competitive disadvantage can be another strong motivator.

By contrast, engineers or condensed-matter physicists in industry seldom feel the same sense of necessity. Lack of experience and exposure to successful applications is one reason. However, it is important to recognize that they often need continuum models and properties. Measurement or correlation is preferred because inaccuracies in quantities or trends from molecular or atomistic models are unacceptable. On the other hand, measurements and correlations are not free of uncertainty, either. Obtaining data is often costly and sometimes impossible. In such situations, when property estimation or mechanistic guidance is necessary, engineers and other technologists find great value in modeling predictions that come with a promise of fundamental soundness. Once they have benefited and accept the limitations and strengths of molecular modeling, they display the same view of modeling as a matter of routine necessity.

Advances in practical applications during the 1990s were remarkable, signaling this transition. The head of corporate research at one chemical company went so far as to say that it was clear that the company's success in applying molecular modeling would be a central basis of its success ten years from now. Not all management feels so strongly, but the present widespread penetration of these methods speaks for their perceived and demonstrated importance.

Impact has been dominated by large companies, both because of the need for specialists and because of hardware and software costs. Some consulting and computational-chemistry companies have sought to provide contract specialists for the needs of smaller companies, finding need and desire but only a limited market to date. This picture is changing as modeling becomes cheaper and as nonspecialist use becomes more widespread and routine, especially among newly graduated employees.

In education and technical communications, both for students and for practicing scientists and technologists, knowledge of molecularly based phenomena and modeling is becoming imperative. Continuum models of materials and of macroscopic behaviors will continue to be important. However, all reaction chemistry occurs at the molecular scale, molecular interactions are the basis of continuum properties, and life sciences are becoming ever more clearly linked to their chemical underpinnings.

Thus we find that molecularly based modeling has now built a record of achievement and limited acceptance. More dramatically, molecular theories and modeling, including electronic-structure theories and modeling, are eventually to become the *lingua franca* for advances in much of science and engineering.

CHAPTER 11

SUMMARY AND VISION

Phillip R. Westmoreland

OVERALL CONCLUSIONS

This report finds strong evidence that in the next ten years molecularly based modeling will profoundly affect how chemistry, biology, and materials physics are understood, communicated, and transformed to technology, both intellectually and in commercial applications. Molecularly based modeling is a basis for communicating, understanding, and developing new technologies like nanotechnology. It creates new ways of thinking—and of achieving.

Thus, to describe successful applications fully, it is not sufficient to describe only the technical features. Just as important are the backgrounds of personnel, the problem or goal's context, approach to the problem, the organization's structure, and the computing infrastructure.

The findings may be grouped into the categories of applications (technical features) and the process of effective application (all aspects of infrastructure). The preceding chapters and the case-study reports in Appendix B give details and evaluation of international activities in this field. Out of these analyses, certain broad findings stand out, and they are summarized here.

MAJOR FINDINGS: APPLICATIONS

(1) Molecular modeling methods have gained acceptance as practical tools in a variety of industries. Tables in Chapter 10 list 55 U.S. companies that have made use of molecularly based modeling, limited to companies making use of modeling specialists—undoubtedly an incomplete list. The tables list an additional 265 companies, universities, or other organizations elsewhere throughout the world. Industry sectors include:

- Pharmaceuticals—the highest-profile area of success and acceptance
- Agricultural chemicals—modeling and informatics for herbicide discovery, similarly to pharma
- Commodity and specialty chemicals including dyes and pigments, oil additives, anti-corrosives, catalysts
- Fuels—resource production, transportation, and processing
- Polymers, glass, and structural materials
- Electronic and photonic materials
- Industrial gases
- Personal-care and food products
- Computer software and hardware

(2) These methods are being adopted because they are proving their value in big ways. Big successes naturally attract the most attention, both within companies and from competitors. The study identified three major success areas as drug discovery, development of homogeneous catalysts, and thermochemistry. Band structure analysis has also had a major effect on condensed-matter physics.

Success has not always been achieved in the way anticipated. Ten and even twenty years ago, "rational drug design" was the path of the future for the pharmaceutical industry. Drugs would be computer-designed based on their chemical and physical bonds ("docking") with molecular regions of cell membranes or other binding sites ("targets"). This has been achieved, and new drugs have been commercialized—though not without the other tools of drug development. However, a rush to combinatorial drug design seemed to make rational drug design outmoded. The new vision was that combinatorial chemistry could be faster and less limited by the imagination. However, trying everything combinatorially is still a daunting task, of course. The present reality is that computer-based, rational exploration of new drug possibilities is increasingly valuable for identifying the library of molecules that can be tested combinatorially in experiments. In this role, and in its pure form, drug design has proven most valuable at narrowing the range of possible candidates.

(3) Many companies now see such modeling as a necessity, while others still consider it a luxury. In the success areas just mentioned, for example, no company would think of carrying out development work without involving molecular modeling. In other areas, modeling is an active area of exploratory research or is expected to be of use. Some potential areas of contribution are in industries that, usually for short-term or long-term market reasons, are in the position of scaling back or of maintaining existing technologies, such as energy companies in the mid-1990s. Other areas of application seem too vast to be impacted by molecular modeling except indirectly, for example cellular biology or atmospheric chemistry or I-beam formation. In others, the technical aspect is minor or absent (economics, advertising, human resources management), so the technical value of modeling is not relevant.

These differences in how modeling is viewed are often based on the idea that investment in modeling must have a direct connection to deliverables. Such an attitude may or may not be sensible. If an investment is intended to achieve a specific deliverable that requires the input of modeling, then the connection is plain. On the other hand, one would not expect that that a new product or a new process would be derived solely from some analytical instrument used in the development process, as crucial as its role may have been. From another point of view, the same would be true for the role of the calculus or of material balances. They are instead key approaches to analyzing and quantifying new technologies, and are so much part of the fabric of development that they should not be singled out.

(4) The principal role of molecular and materials modeling in the chemical industry is to speed product development and guide experiment. Energetic participation of modelers in the early stages of process or product development has repeatedly proven to be the strongest, most effective model for problem solving. As several hosts put it, the greatest value can be using simple calculations to reduce the number of possible alternatives. In product design and the early stages of process design, accurate trends are often more valuable than accurate numbers. Modeling can guide choices and, even more often, can identify productive avenues for experiments. In contrast, the technical-services model of providing molecularly based modeling has proven quite unsuccessful in when it is not coupled closely into internal development teams or the internal clients.

(5) Science and methodology have made great strides, and successful applications encourage more. Applications benefit from scientific advances, but they also spur new science. Scale-bridging is an excellent example, both in time and in spatial size. The small groups of atoms that could be analyzed by quantum chemistry methods—initially no more than H and H_2—compared to the many-atom domain of most applications has stimulated numerous scientific advances:

- Molecular simulation using Newtonian molecular dynamics and Monte Carlo stochastically generated system changes
- Semi-empirical molecular-orbital methods

- The hierarchy of quantum chemistry methods that provide a time-accuracy trade-off, including method/basis-set extrapolation methods like G1/G2/G3 and CBS, and domain-substitution methods like ONIOM
- Electronic density functional theory, formulated in the physics community and used to begin modeling condensed matter
- *Ab initio* molecular dynamics, notably starting with the Car-Parrinello formulation of on-the-fly DFT calculations
- Nested and domain-coupled simultaneous methods like the effective medium theory of Nørskov, the embedded-atom or cluster methods, the FE/MD/QM approach of Abraham et al., and the "seamless zooming" of the Doi Project in Japan
- Periodic boundary condition models that capture longer-range influences
- The general class of mesoscale modeling, often using continuum equations to model supermolecular but subcontinuum domains

At the same time, important new applications have reached the nanometer scale, smaller than continuum methods can reach, where molecularly based modeling is the tool of necessity. Likewise, the chemical knowledge of the genome (set of genes existing in the structure of DNA) requires knowledge of how instructions are transformed into the proteome (set of proteins that exist and carry out function). Modeling has an important role in deciphering this process and in discovering how a given strand of amino acids can fold itself into the protein geometry that has a particular function. However, even the rapid time scale of protein folding is presently a long simulation time—perhaps a year on the IBM "Blue Gene" dedicated supercomputer that is under development. As a closely linked topic, computer-based informatics is also necessary to exploit the wealth of data and computed results, driving developments in information science.

(6) Molecular modeling is being used to complement and interpret analytical measurements. In some cases, predictions now have less uncertainty than the calorimetric measurements, effectively replacing them. Modeling has become an invaluable part of optical spectroscopy, NMR, and crystallography. Future roles will include interpretation of gas chromatography, vibrational circular dichroism, and mass spectrometry, based on already successful calculations.

(7) Striking similarities exist among needs of different industrial sectors. Methods often cut across sector and disciplinary boundaries, e.g., scale-bridging methods, models for solvation and real mixtures, and models that match chemicals to the biochemical basis of a biological behavior, the basis for drug design but also for predictive chemical toxicology. With predictive toxicology, chemical products could be designed that have a desired function but with less hazard than alternatives.

(8) The top needs required by industry are methods that are "bigger, better, faster," more extensive validation, and multiscale techniques. Multiscale modeling is often built on models that apply in a limited scale of size and time, such as a single-point energy calculation on an ideal gas (isolated) molecule at 0 K. Modeling at the restricted scales can often be coupled to continuum properties by statistical mechanics (e.g., to ideal gas thermochemistry). There is a pressing demand to generate all these results faster and more accurately. To that end, testing of methods is essential to establish the range of validity. Note that this concept of "validation" is distinct from the mistaken illusion that such tests would establish that a method is valid at all conditions or scales.

(9) Looking ahead, molecularly based modeling is the key to modeling a vital technology for the future: applying nanostructure and nanoscale devices. Nanotechology is the technology of nanometer-scale materials and processes. A typical atom has a van der Waals diameter of a few tenths of a nanometer, so molecules and macromolecules are nanometer-sized and larger. All reactions and many macroscopic properties have their origins at this scale, linked to the macroscale by statistical mechanics and solid-state physics. At the same time, the properties may be dominated by electronic structure, non-bonding interactions, or meso-scale, supermolecular behaviors. Each of these domains is the province of some type of molecular modeling, and it is increasingly possible to couple these domains to each other and the macroscale.

MAJOR FINDINGS: PROCESS OF EFFECTIVE APPLICATION

(1) The principal role of molecular and materials modeling in industry is to speed product development and guide experiment. Most commonly and effectively, this is by developing mechanistic hypotheses and a rational, strategic approach to problem solving. Beyond the discovery stage, modeling can aid relative ranking of properties for molecules, materials, and formulations. At the process design and operation stage, generation of chemical and material properties is most important. A surprising role is the use of modeling in defining and defending intellectual property. It is not enough for specialists to understand these roles, though; it is critical that managers understand the power and limitations of these approaches. Modelers bear the responsibility of helping insure that expectations are not just visionary but also reasonable.

(2) Eliminating dead ends can be especially valuable. Assigning value to the path not taken is usually difficult, yet this role of modeling can make the most difference financially. Long-term, the best way to evaluate this impact is by the increase in productivity. The product pipeline of pharmaceutical companies has to be full, and modeling has demonstrated its value at every company by increasing the proportion of promising and fruitful leads. Likewise, this screening role has proven its value in development of homogeneous catalysts, of reaction mechanisms, of polymers with particular properties.

(3) Failure to couple modeling sufficiently with experiments can lead to frustration, delays, and irrelevance. Coupling modeling with data can be powerful. Not doing so can be disastrous. Nevertheless, failure to communicate this requirement for coupling between experiment and modeling is a major cause of unsuccessful modeling activity. This problem can be attributed in part to organizational structures and geographies that impair communication between the people with the models and the people with the data.

(4) A healthy division has developed among specialists, non-specialist users, and internal clients. Not everyone needs to be a modeler. However, this has been a frequently stated goal of many newly formed groups of modeling specialists. In more established centers of activity, it has become clear that many scientists and engineers are benefiting from the three-dimensional display and manipulation of molecular information. More and more, new employees are familiar with such basic tools, and they expect to have access to the necessary hardware and resources—and specialists to handle more sophisticated problems.

The most cost-effective use of modeling specialists is to address problems that require significant sophistication in modeling. That includes helping identify which problems of their clients require such high-level consideration. Educating internal clients and managers, another common role, is then seen not as training to be modelers but helping people become savvy about what problems can be addressed, what levels of accuracy can be achieved, and what resources are available.

(5) Using molecular modeling for intellectual property should be understood as extending beyond discovery. Discovery and invention are obvious roles for modeling. However, defining and defending intellectual property is a broader idea. Modeling has had other important impacts in intellectual property, including broadening of patent claims, development of offensive patents ("foot in the door"), and evidential use in court proceedings.

(6) Computing platforms are shifting toward commodity processors, even for much high-performance computing. Bigger and faster calculations require higher-capacity, faster computers ("high-performance computing"). The computers required for fruitful molecular modeling are smaller than ever. Mainframes were augmented in the 1970s by minicomputers. This blend evolved to a mix of vector supercomputers and workstations for serious calculations, although specialized parallel computers proved of some value and PCs quickly became useful for routine office tasks, graphical display, and handling of modest databases.

At present, PCs have achieved workstation capability for large but now routine calculations, moving toward UNIX-like operating systems (often hidden under effective user interfaces) and outperforming UNIX workstations only a few years older. For mammoth calculations, networked vector supercomputers, workstations, and especially commodity PC microprocessors have become the bases for parallel supercomputers of teraflop power. The Beowulf concept, attributed to Thomas Sterling at NASA in 1993, was based on physically connecting PCs with fast Ethernet switches and operating them as a parallel computer using the LINUX operating system. Many modelers in companies, government, and academics are

all moving to Beowulf clusters to exploit parallel computing. However, we note that many industrial modelers expressed reluctance to get involved with this approach due to the problems of installation and system management.

(7) Software development must avoid becoming narrowly focused on the drug-discovery problem. The financial incentive varies for drug development, tied as it is to government licensing and reimbursement policy. Nevertheless, the potential is reflected in the book title "The Billion Dollar Molecule: One Company's Quest for the Perfect Drug" (Werth 1995), which is based on events at Vertex Pharmaceuticals, including its use of computer modeling for drug design and interpretation of protein crystallography. With the acquisition of MSI, the largest molecular-modeling software company, by Pharmacopeia, a drug discovery and discovery-methods company, many modelers involved in materials and chemicals applications became concerned that they would be shut out. That has not proven to be the case, as evidenced by subsequent events and the MSI and Pharmacopeia reports in Appendix B.3. However, other companies completely abandoned the non-bio market for molecular modeling software. The science of molecular modeling and the technologies of its software and applications are too strong and too valuable in the long term to be abandoned for long. One example was the movement of Tripos back into materials modeling with polymer QSPR codes.

(8) Technology-oriented modelers and modeling-savvy technologists are needed. Specialist and non-specialist modelers both need a productive combination of expertise and attitude. This combination is easiest to achieve in the workplace, where the need for this blend is always apparent. The need can be harder to recognize in a corporate research environment, where despite economic pressures, some companies seek to preserve a "Bell Labs" environment of blue-sky fundamental research. (Note that Bell Labs was founded for applied research, and it has always maintained a blend of applied and fundamental research.) Expertise is easiest to teach in an educational institution, and it is important to teach collaborative attitudes as part of science education.

VISION

A number of projected expectations for the future follow directly from the findings described above. More companies will find more use for these methods as more of their personnel have experience with the methods and results. Molecular modeling results or calculations will become an intimate part of analytical chemistry, embedded in the software that runs instruments and interprets results. Integrated multiscale modeling will be used more and more heavily, and more and more predicted results will be relied on, methods having been found to give results as accurately as certain experiments.

Some broad visions should be expressed as well. The following sections discuss (1) the aid of modeling in describing molecular bases of chemical, physical and biological science, (2) the impact of great increases in computing power, (3) the potential role of expert systems, and (4) the value of improved organizational strategies.

A Scientific Common Language

Molecular theories and modeling, including electronic-structure theories and modeling, are likely to become the *lingua franca* for advances in much of science and engineering.

Chemistry, physics, and biology originally were all based on observation and manipulation of human-scale phenomena. They were empirically based and relied heavily on classification and analogy. Engineering shared these origins, adding economic considerations and the quantitative component of physics.

Continuum physics and speculations on the nature of electrons in atoms in the 1800s foreshadowed the developments of statistical mechanics and quantum mechanics. The advent of physical chemistry and of concepts of molecular structures in the late 1800s similarly foreshadowed understanding of the chemical bond that was developed in the 1930s and quantum chemistry methodologies that were developed in the 1950s. In biology, classification led logically to Mendel's empirical model of genetics. In the U.S., the stimulation of research on molecular biology by the Rockefeller Foundation in the 1930s led to Watson and

Crick's discovery of the chemical-physical structure of DNA in 1953 (Kay 1992) and to subsequent understanding of the chemical basis of cellular features and functions. Engineering embraced continuum models of mathematical physics beginning in the late 1940s and the 1950s, but application of quantitative atomic and molecular understanding of chemistry, atomistic physics, and biochemistry is only now achieving real impact.

Many of the advances of the 2000s will be underpinned by quantitative understanding of the phenomena of chemistry, physics, and molecular biology. Molecularly based modeling is a central method by which quantitative understanding is being achieved, even to the point of prediction. Beyond understanding is sharing that understanding. The development of products and processes changes when the team can reflect on the impact of molecular properties.

Education in the classroom takes on a different cast when observable properties are related to visually or numerically observable properties of molecules and ensembles of atoms and molecules. Public awareness is affected as well; even if unclear on the details, the general public is familiar with concept of DNA as a helical molecule that carries genetic information relevant to reproduction, genetic disease, and even identification more accurate than fingerprints. Molecular modeling then becomes part of the public's unseen scientific infrastructure, appearing without fanfare as sketches of molecules in newspapers and advertising, and embedded in the creation of new products.

Unimaginable Computational Power Will Lead to Addressing Unthinkable Problems

Advances in the power of computer hardware, pulling along with it advances in software, will dramatically change the ways of thinking and modeling in chemistry, physics, and biology. Many vastly more complicated problems will become tractable, especially those that scale linearly with the number of particles, atoms, or electrons. At the same time, the ability to handle more realistic details will demand new fundamental understanding.

Consider the example of modeling a simple gas-phase flame, burning a premixed gas flowing in smooth laminar streamlines. Modeling the full chemistry of such a flame, computationally predicting the combination physics and elementary chemical reactions that give rise to it, did not happen until the late 1970s (Warnatz 1981). The first model solved the steady-state equation for conservation of mass in the presence of convection, thermal and molecular diffusion (modeled with pseudo-binary Fickian diffusivities), and reaction. The calculations required a temperature profile in lieu of solving an energy equation, Lennard-Jones parameters for transport properties, a set of about 60 reactions and 30 chemical species, Arrhenius and non-Arrhenius rate constants, and polynomial-based expressions for the ideal-gas thermochemistry of each species. Agreement with overall parameters like flame speeds was very good, and later work extended the evaluation to spatial profiles of the molecules and radicals and to much larger mechanisms (Westmoreland et al. 1986).

For several reasons, these achievements were not previously possible. It had been nearly inconceivable to solve the stiff, highly nonlinear differential equations for 60 reactions simultaneously, even on a one-dimensional grid. However, the major reasons were insufficient computing hardware power and lack of suitable numerical methods for solving stiff ODEs. When computing power advanced sufficiently, the inconceivable became possible.

At the same time, the reality is that rate constants and thermochemistry and transport properties are statistical representations of molecular-level variation. The physics is described well as a continuum of ideal gas, although more realistic diffusion models yield real improvements. Many of the reactions are simple abstractions of H atoms by free radicals, described well by transition state theory using parameters from computational quantum chemistry. Many of the reactions—including all of the oxidation and pollutant-forming growth reactions, association and dissociation reactions that are collisionally mediated. A simple thermal dissociation of a molecule, dilute in an inert gas like argon, can be described well by rovibrationally quantized RRKM theory and *ab initio* parameters.

However, conditions can arise when a reaction rate constant is insufficient, even including dependence on temperature, pressure, and third-body concentration. An association reaction, like OH addition to ethylene,

forms a chemically activated radical; i.e., it contains the absolute thermal energy of the reactants plus the energy of the formed bond, now released into the radical as rovibrational energy. Collisions can lower its energy into the thermal energy distribution, or internal rearrangement or decomposition may dissipate the excess energy. How is collisional stabilization affected if the collider can react with it? How is the population affected by other channels forming the adduct radical, such as H + vinyl alcohol? What can we do about electronically excited states of adduct and collider? Returning to the central point, how can molecular modeling describe such complications of intra- and intermolecular interactions?

The only apparent way to model all these processes is to model the microcanonical states of the species. What is the microcanonical rate constant for reaction of the chemically activated adduct with a colliding radical or molecule? How is transport affected? It would be almost unimaginably complicated to attempt to solve such a problem, much less in a turbulent flow. Nevertheless, the first attempts at addressing such problems have already been made.

Dramatic increases in computer power will permit these and other seemingly intractable calculations within the next ten years. The Blue Gene protein-folding project will use a special-purpose petaflop computer (10^{15} floating-point operations per second), two orders of magnitude faster than ASCI White, the fastest present computer. This petaflop computer is expected to be operational and ready for its first year-long calculation in about 2005. Projected computer power available as soon as 2010 would allow such huge calculations to be performed two or three orders of magnitude faster still—a day-long calculation instead of a year.

Will Artificial Intelligence Ever Replace the Real Thing?

The value of computed results will be greatly enhanced when they can be generated near-instantly, but they become invaluable when their meaning is clearly linked to the application's or developer's need.

Advances in computational power will advance our understanding and our ability to apply computational chemistry, physics, and biology. Computed results already can yield empirical relationships like QSARs and QSPRs based on correlation of experimental properties with computed electronic or molecular properties. It is a modest step to empirical correlations of accepted, computed properties with computed molecular characteristics. Just as great advances in technology often result from experimental observation, with understanding following later, molecularly based modeling is increasing in accuracy and complexity such that it may yield useful results ahead of our understanding the results.

However, chemical, physical, and biological understanding is still likely to guide such calculations. Dr. Bruce Eichinger of MSI / Pharmacopeia summed it up as follows:

> Usually, chemical problems don't get solved by just computing something. One needs a clearly defined strategy based on clear intuition. You have to solve 90% of the problem in your head and then do the calculation to test various hypotheses.

Effective modeling and management of its results is a more likely role for artificial intelligence or expert systems. There is much art in the successful operation of molecular modeling codes. Likewise, the wealth of computed results already demands intelligent use of visualization to capture essential features. There may or may not be a single number or equation at the end of a calculation. Managing the flood of information demands intelligent prescreening and data mining.

The choices of theoretical method and basis set (quantum chemistry) or interaction potential (molecular simulation) make up the lowest level of decision-making. Which combination should be selected for the optimal balance of time and accuracy? Would a quick calculation yield a result with adequate accuracy? In a product or process development environment, such a result would be far more valuable than a longer, more accurate calculation. On the other hand, the more accurate number may be required for design of a product or process that must have long life, reliable performance, or large investment. In addition to using the right theory, there is the matter of coaxing the code to give the right results most efficiently, selecting the best settings for the numerical method.

Codifying this information into a user interface would be approximate but useful. Imagine specifying the desired property or reaction, an estimate for the configuration of the molecule or transition state or ensemble,

the confidence required, and the desired time. A suitable interface could evaluate the request and return a recommendation of method with estimates of time and other resources required, executing it if acceptable.

Likewise, a user interface could screen the results, delivering the requested information but also identifying relevant or unusual features that came out in the computation. A selection of ways to interrogate the results or extend them could be offered. Such advances would increase reliability as well as efficiency. The challenge has long been to make codes foolproof and user-friendly. Dr. Ken Cox of Rice University, then at Shell R&D in Houston, noted that it is increasingly important to make them "fool-friendly and user-proof."

Organizational Strategies Can Be Improved

Experimentalists, process technologists, designers, and management not only will become used to results of molecular modeling but will expect them. The most important components of operational strategy are then (1) close coupling of the modelers to development teams or modeling clients at the earliest possible stage and (2) shared understanding of what modeling can deliver within time, computing-resource, and budget constraints.

To repeat an observation made earlier in this chapter, the principal role of molecular and materials modeling in industry is to speed product development and guide experiment. Modeling rarely points to new substances or processes by itself. Modeling is basically a tool for analyzing hypotheses and for predicting properties for defined conditions.

The most productive time for generating and testing hypotheses is in the early stages of development. Dow modelers described an internal meeting held between modelers and polymerization chemists on developing homogeneous catalysts. To address questions and proposals made early in the meeting, an SGI workstation was brought in and an impromptu session of calculations was held to test hypotheses as they were offered. These were simple calculations with Spartan software and 3-D glasses. Successful patent filings for new catalysts came out of the meeting.

Second, realistic expectations are important. As with any specialized information, modeling can seem like a black box to non-specialists. At the extremes, observers can accept the results without question, or they can skeptically reject results outright. Either response is bad, especially when the respondent is a client or manager. Molecular visualization can exacerbate the first problem because a picture of molecules or interaction implies strongly that the result is a firm reality. Failure to meet expectations, even if the expectations were unrealistic, damages credibility of the modelers and also of modeling.

Thus, communicating the limits of modeling is crucial to its successful use. A key part of the "validation" process described earlier is to define the limits of accuracy and to identify calculations that give distinctive errors. This is an open-ended task for the future, but some software vendors give benchmark comparisons of time, resources, and accuracy for different methods. The Computational Chemistry Group of NIST (http://www.nist.gov/compchem/) has made another significant attempt in this area, establishing a Web-based "Computational Chemistry Comparison and Benchmark Database" and a "Sicklist Database."

The requirements for success are still the right tools being used by the right people to solve the right problems. Dramatic advances in computer hardware and modeling software—the tools—are expected to continue, making more successes possible and making previously unapproachable tasks approachable. Computer power is not enough by itself, though, and it is unlikely ever to be. The right people and the right problems—which ultimately comes down to the right people again—are the real secrets to success.

The key challenges for future success in molecular modeling, therefore, are for academic institutions to develop the right people and for industrial organizations to adopt wise strategies. In addition, governments around the world have proven effective at helping meet this challenge by supporting quality education and by stimulating collaborations between academia and industry

REFERENCES

Kay, L.E. 1992. *The Molecular Vision of Life: Caltech, the Rockefeller Foundation and the Rise of the New Biology.* Oxford University Press.

Warnatz, J. 1981. *Proc. of the Combustion Institute*, 18, 369.

Werth, B. 1995. *The Billion Dollar Molecule: One Company's Quest for the Perfect Drug.* Touchstone Books.

Westmoreland, P.R., J.B. Howard, and J.P. Longwell. 1986. *Proc. of the Combustion Institute*, 21, 773.

APPENDICES

APPENDIX A. BIOGRAPHICAL SKETCHES OF PANEL MEMBERS AND OTHER TRAVELING-TEAM MEMBERS

PANEL MEMBERS

Professor Phillip R. Westmoreland (Panel Chair)

Department of Chemical Engineering, University of Massachusetts Amherst
Address: 159 Goessmann, Box 33110, Amherst, MA 01003-3110
Email: westm@ecs.umass.edu
URL: http://www.ecs.umass.edu/che/westmoreland.html

Phillip R. Westmoreland is Professor of Chemical Engineering at the University of Massachusetts Amherst. He received a BSChE from North Carolina State University (1973), an MSChE from Louisiana State University (1974), and a PhD from MIT (1986). From 1974 to 1979, he was a research engineer at Oak Ridge National Laboratory, conducting coal conversion research.

His present research focuses on applied kinetics of gas-phase combustion, polymer flammability, plasma-enhanced chemical vapor deposition, and homogeneous catalysis, obtained both from experiments and from computational chemistry. He is the author of fifty publications, 116 presented papers and posters, and an edited book.

Professor Westmoreland is active in the American Chemical Society, the American Institute of Chemical Engineers, and the Combustion Institute. He organized a 1991 ACS fuel chemistry symposium on "Combustion Chemistry," 1994 and 1996 AIChE Symposia on "Computational Chemistry and Its Industrial Applications," AIChE's discussion forum in this area, and was and is co-chair of 1998 and 2001 AIChE Topical Conferences on "Applying Molecular Modeling and Simulation." In 2000, he was co-chair of a CACHE-sponsored first conference on "Foundations of Molecular Modeling and Simulation."

In AIChE, he is the founding chair of the Computational Molecular Science and Engineering Forum (2000-2001). He has served on the boards of the *I&EC Research*, *International Journal of Chemical Kinetics*, Eastern States Section of the Combustion Institute, and the CACHE Corporation, and he is chair-elect and secretary of the latter two organizations, respectively.

Awards include NSF Presidential Young Investigator (1990-95), AIChE Public Relations Award (1977), GE Outstanding Teaching Award (Univ. Massachusetts, 1990), Outstanding Junior Faculty Award (Univ. Massachusetts, College of Engineering, 1990), and BCR/R.A. Glenn Award for Best Paper (ACS Fuel Chemistry Div., 1992).

Professor Peter A. Kollman (Panel Vice Chair; b. 1944, d. 2001)

Department of Pharmaceutical Chemistry, University of California San Francisco
URL: http://www.amber.ucsf.edu/kollman.html

Peter Kollman received his BA from Grinnell College and his PhD degree from Princeton University (1970). He was a NATO Postdoctoral fellow at Cambridge University in England 1970-71. He was on the faculty of the Dept. of Pharmaceutical Chemistry, School of Pharmacy at the University of San Francisco starting in 1971, where was most recently Professor of Chemistry and Pharmaceutical Chemistry. He was awarded a Career Development Award from the NIH, 1974-79, and in 1995 won the American Chemical Society Computers in Chemistry Award. He published over 300 papers on applying computational methods to molecules of organic and biological interest and has been the mentor of over 50 graduate and postdoctoral students.

Anne M. Chaka, PhD

Group Leader, Computational Chemistry Group, National Institute of Standards and Technology
Address: Physical and Chemical Properties Division (838), Chemistry Building (222), Room A155, 100 Bureau Drive, Stop 8380, Gaithersburg, MD 20899-8380
Email: anne.chaka@nist.gov
URL: http://www.nist.gov/compchem/

Currently at the National Institute of Standards and Technology (as of the publication date of this report), Anne Chaka was at the Lubrizol Corporation for most of the duration of the study, where she was in charge of the computational chemistry and physics program in R&D. Research interests include:

- Atomistic descriptions of corrosion, wear, and lubrication mechanisms
- Surface passivation
- Pericyclic reaction mechanisms
- Free radical chemistry
- Heterogeneous and homogeneous catalysis
- Thermochemistry
- Combustion and oxidation

In 1999-2000, she was on an industrial sabbatical from Lubrizol with a Max-Planck-Society Fellowship at the Fritz-Haber-Institut in Berlin, Germany with Matthias Scheffler. Previously, she has been Technical Director of ICN Biomedicals, Inc., an analytical research chemist for Ferro Corporation, and a Cray programming consultant to Case Western Reserve University for the Ohio Supercomputer Center.

Present and past professional positions include the Executive Committee of the Physical Chemistry Division, secretary of its Theoretical Chemistry Subdivision, American Chemical Society, the editorial board of *Molecular Simulation*, Physical Committee Chair and Alternate Councilor of the ACS Cleveland Section, ACS Morely Award Committee, ACS Industrial Award Search Committee, and the Anna Louise Hoffman Award Committee.

Educational background includes a PhD in theoretical chemistry from Case Western Reserve University, an MS in clinical chemistry from Cleveland State University, and a BA with Honors in chemistry from Oberlin College.

Professor Peter T. Cummings

Department of Chemical Engineering, University of Tennessee
Oak Ridge National Laboratory
Address: University of Tennessee, Knoxville, TN 37996-2200
Email: ptc@utk.edu
URL: http://www.che.utk.edu/grad/brochure/ptc.html

Peter T. Cummings received his PhD in applied mathematics (statistical mechanics) from the University of Melbourne (Australia) in 1980. After post-doctoral appointments in physics at the University of Guelph (1980) and in chemistry and mechanical engineering at the State University of New York at Stony Brook (1981-83), he joined the Department of Chemical Engineering at the University of Virginia as an assistant professor in 1983, attaining the rank of full professor in 1991. In January 1994, he assumed his current joint position as Distinguished Professor in the Departments of Chemical Engineering, Chemistry and Computer Science at the University of Tennessee and Distinguished Scientist in the Chemical Technology Division at Oak Ridge National Laboratory.

Professor Cummings is the recipient of many awards, including a Camille and Henry Dreyfus Foundation Research Award for Newly Appointed Faculty in the Chemical Sciences in 1983, the inaugural Gold Medal for Professional Excellence awarded by the Convocation (alumni association) of the University of Newcastle in 1988, and the 1998 Alpha Chi Sigma Award from the American Institute of Chemical Engineers (its premier research award recognizing research achievement over the past decade). Since 1991, he has been a

member of the editorial boards of two international statistical mechanics journals, *Molecular Physics* and *Fluid Phase Equilibria*; in 1997, he was elected to the Editorial Advisory Board of *Molecular Physics*. He is a member of the American Physical Society, American Institute of Chemical Engineers, American Chemical Society, Sigma Xi, Society of Rheology, and Materials Research Society.

Professor Cummings is internationally renowned for his research on molecular approaches to predicting physical properties in systems of industrial interest, such as supercritical aqueous solutions, alkane fluids, and polymer solutions. He is at the forefront in the application of massively parallel supercomputers to predicting thermophysical properties. Professor Cummings is the author of over 190 refereed publications, a frequent invited speaker at international conferences and a consultant to several companies. His research is supported by the National Science Foundation, the Department of Energy (Chemical Sciences Division and Environmental Management Science Program), Lockheed-Martin Energy Research Corporation, and the Chemical Technology and Chemistry and Analytical Sciences Divisions of Oak Ridge National Laboratory.

Professor Keiji Morokuma

Department of Chemistry, Emory University
Address: 1515 Pierce Drive, Atlanta, Georgia 30322
Email: morokuma@emory.edu
URL: http://fserver.chem.emory.edu/users/groups/morokuma.html

Keiji Morokuma is the William Henry Emerson Professor of Chemistry and Director of the Cherry L. Emerson Center for Scientific Computation at Emory University in Atlanta, Georgia, where he has taught since 1993. He received his PhD degree from Kyoto University under Kenichi Fukui, and was a post-doc with Martin Karplus at Columbia and Harvard. He was Professor of Chemistry at the University of Rochester from 1967-1977. From 1977 through 1992, he served as the founding director and professor of the Department of Theoretical Studies and the Computer Center of the Institute of Molecular Science (IMS) in Okazaki, Japan. He is recipient of the Annual Award of the International Academy of Quantum Molecular Sciences, the Award of the Chemical Society of Japan, and the Schrödinger Medal of the World Association of the Theoretical Organic Chemists. He has published more than 420 scientific papers on methods and applications of quantum chemistry.

Matthew Neurock, PhD

Department of Chemical Engineering,University of Virginia
Address: 102 Engineers' Way, P.O. Box 400741, Charlottesville VA 22904-4741
Email: mn4n@virginia.edu
URL: http://www.che.virginia.edu/neurock.html

Matt Neurock is Associate Professor of Chemical Engineering at the University of Virginia. Prior to coming there, he received his PhD degree from the University of Delaware under Michael T. Klein, was a post-doctoral associate with Rutger van Santen at the Technical University of Eindhoven, and worked in applied computational chemistry at E.I. du Pontâs Engineering Experiment Station in Wilmington, Delaware. He is widely recognized for his work in modeling of catalysis using quantum chemistry and kinetic Monte Carlo methods to examine the properties and performance for a wide range of different materials including metals, bimetallics, metal oxides and zeolites for their use as heterogeneous catalysts, catalytic electrodes for fuel cells, and magnetic materials for memory device fabrication. He has been presented several young faculty awards, including from Ford, DuPont, and the National Science Foundation (CAREER Award).

Ellen B. Stechel, PhD

Ford Motor Company
Address: Scientific Research Laboratory – MD3083, 2101 Village Rd, Dearborn, MI 48124

Ellen B. Stechel has been Manager of the Chemistry & Environmental Science Department, Ford Research Laboratory, Ford Motor Company. She is currently on temporary special assignment outside of Research in Core and Powertrain Engineering, within Ford Motor Company. She received her AB in mathematics and chemistry from Oberlin College in 1974. She received an MS in physical chemistry and completed her PhD

in chemical physics from the University of Chicago in 1976 and 1978, respectively. After a postdoctoral research position at the University of California, Los Angeles, she joined Sandia National Laboratories in a technical staff position in Condensed Matter Physics in September 1981.She has a broad range of management, policy, and technical experience in the fields of science policy, computational science, materials science, and environmental science. Her research interests have been varied, including computational theory of reactive scattering of A+BC and electron transfer reactions; quantum chaos; electronically stimulated processes on surfaces; electronic structure of high T_c cuprates; and advanced algorithmic development in density functional theory. In 1993 she joined the Advanced Materials and Device Sciences Department at Sandia National Laboratories and became the manager of that department in 1994. She continued in the management position until November 1998. From its inception in December 1997 until she left Sandia National Laboratories in November 1998, she was co-technical director of a virtual center, Center for the Simulation of Complex Materials. She is now serving as a co-coordinator for the Computational Materials Science Network, a virtual network, sponsored by the Department of Energy, Office of Science, Basic Energy Sciences, Division of Materials Science.

Dr. Stechel has served as a Senior Editor for the *Journal of Physical Chemistry* and in numerous professional society positions including Board of Directors, American Vacuum Society; Panel on Public Affairs, American Physical Society; Program Chair, Chair and Past Chair of the Division of Physical Chemistry, American Chemical Society. The positions that she currently holds in the American Chemical Society are Councilor, Committee on Science, and Corporate Associates. She also serves on the Advanced Scientific Computing Federal Advisory Committee for the Department of Energy and on the Chemical Sciences Roundtable, NRC.

Professor Priya Vashishta

Concurrent Computing Laboratory for Materials Simulations
Department of Physics & Astronomy and Department of Computer Science
Louisiana State University
Address: Nicholson Hall, Physics Department, Louisiana State University, Baton Rouge, LA 70803-4001
Email: priyav@bit.csc.lsu.edu
URL: http://www.cclms.lsu.edu/cclms/group/Faculty/vashishta.html

Priya Vashishta is Cray Research Professor of Computational Science at Louisiana State University. He is a founding member and the Director of the Concurrent Computing Laboratory for Materials Simulations. He has appointments in the Department of Physics & Astronomy and the Department of Computer Science. Prior to LSU, he was at Argonne National Laboratory, where he was Senior Scientist from 1984 and the Director of the Solid State Science Division from 1979 to 1982. He received his PhD from Indian Institute of Technology, Kanpur in 1967. He was a visiting professor/scientist at IBM Research Center (1972), Bell Labs (1976), University of California, San Diego (1976-77), Institute of Theoretical Physics, Santa Barbara (1983), and Cornell University (1986-87).

Professor Vashishta has been involved in various aspects of materials simulations for over 25 years. His current research interests include very large-scale atomistic simulations of novel materials and electronic devices on massively parallel and distributed computers. These include: computational algorithms for multi-million particle materials simulations on parallel architectures; molecular dynamics simulations of structural and dynamical correlations in covalent glasses, ceramics, and nanophase composites; simulation of quantum processes in nanoscale devices; sintering of nanophase ceramics; crack propagation and fracture in amorphous and crystalline ceramic films and coatings.

The Concurrent Computing Laboratory for Materials Simulations (CCLMS) has two main objectives: (1) multidisciplinary research involving grand challenges in materials simulations; and (2) the development of undergraduate and graduate education in computational sciences. The CCLMS has two parallel computing laboratories. It has also created a dual-degree program consisting of a PhD in physical sciences and an MS in computer science.

Professor Vashishta has published more than 230 papers, given 110 invited talks at conferences and workshops, and presented 90 seminar talks at universities. He has edited 10 books and is on the editorial

boards of *Journal of Computational Materials Science* and *Journal of Plasma and Ions*. He has served on a number of committees and panels of NSF and DOE.

He has received the University of Chicago Award for Distinguished Performance at Argonne National Laboratory (1976); Japan Society for the Promotion of Science Senior Fellowship Award (1985 and 1989); Brazilian Science Research Council Fellowship Award (1985); United Nations Development Program Fellowship Award (1990); and the DARPA Award for ULTRA Electronics Sustained Excellence (1997).

OTHER TRAVELING TEAM MEMBERS

Professor Sharon C. Glotzer

Departments of Chemical Engineering, Physics, and Materials Science and Engineering; University of Michigan
At the time of the study: Director, Center for Theoretical and Computational Materials Science, and Physicist, Polymers Division, National Institute of Standards and Technology
Address: Department of Chemical Engineering, Room 3014
H.H. Dow Building, 2300 Hayward; Ann Arbor, MI 48109-2136 USA
Email: sglotzer@umich.edu

Sharon Glotzer joined NIST in 1993 as an NRC postdoctoral fellow in the Polymers Division, after receiving her PhD in theoretical condensed matter physics from Professor H. Eugene Stanley's group at Boston University. She became a permanent member of the Polymers Division in the Materials Science and Engineering Laboratory in 1995. Dr. Glotzer is the director and co-founder of the NIST Center for Theoretical and Computational Materials Science. In 2001, she joined the Department of Chemical Engineering at the University of Michigan as an Associate Professor with joint appointments in Physics and in Materials Science and Engineering.

Dr. Glotzer has published close to 40 scientific papers in refereed journals, edited several conference proceedings, and organized or co-organized a dozen workshops and scientific meetings. She has given close to 30 invited talks at national and international conferences on her research on theory and simulation of soft materials, and on the growing field of computational materials science. Dr. Glotzer is a member of the American Physical Society, Materials Research Society, American Chemical Society, and AICHE. She is currently a member of the nominating committee for the APS Division of Computational Physics, and a member of the judges panel for the APS Metropolis Award. Dr. Glotzer serves on several scientific advisory and editorial boards, including the NSF Science and Technology Center for Advanced Liquid Crystalline Optical Materials and *PhysChemComm*. She has been a founding member of the Science Advisory Committee of the Los Alamos National Laboratory Institute for Complex Adaptive Matter. She serves as a PhD advisor to students in the Chemical Physics Program at the University of Maryland, and supervises a number of postdocs in her group.

Dr. Glotzer has received a number of awards, including an NRC Postdoctoral Fellowship Award from 1993-1995; a Senior Visiting Fellowship Award at the Centre for Chemical Physics, University of Western Ontario in May 1996 for research related to pattern formation in complex fluids, and frustration in glass-forming systems; the Department of Commerce Bronze medal award for co-founding the CTCMS in 1994; and the 1998 Presidential Early Career Award for Scientists and Engineers in recognition of her broad contributions to the theory of complex fluids and soft materials, and for her leadership in the emerging field of computational materials science.

Dr. Karl K. Irikura (European site visits)

National Institute of Standards and Technology
Address: Physical and Chemical Properties Div., MS 8380, 100 Bureau Drive,
Gaithersburg, MD 20899-8380
Email: karl.irikura@nist.gov
URL: http://www.nist.gov/compchem/irikura/irikura.html

Karl Irikura received an AB in chemistry from Harvard College in 1984 (M. Karplus; biomolecular modeling) and a PhD in physical chemistry from the California Institute of Technology in 1991 (J. L. Beauchamp; experimental ion chemistry). In 1991 he did postdoctoral work at Caltech (W.A. Goddard; *ab initio* organometallic chemistry) before moving to NIST as a National Research Council postdoctoral associate (J.W. Hudgens; experimental resonance-enhanced multiphoton ionization spectroscopy). In 1993 he became a permanent staff member at NIST and is now in the Computational Chemistry group there. His current research focuses on computational thermochemistry, with additional collaborations in molecular spectroscopy and the prediction of total ionization cross sections. He has authored or coauthored about 45 papers and has edited one book, *Computational Thermochemistry* (with D. J. Frurip of Dow Chemical Co.). He was recently presented the Young Scientist Award of the NIST chapter of Sigma Xi.

Dr. Randall S. Jones (Japanese site visits)

Physics Department, Loyola College in Maryland
Address: 4501 North Charles Street, Baltimore, MD 21210
Email: rsj@loyola.edu

Randall Jones received his doctorate in theoretical condensed-matter physics from Cornell University in 1983. He worked as a post-doctorate associate for three years at the Quantum Theory Project at the University of Florida, before accepting his current position in the physics department at Loyola College in Maryland. His research interests are centered around computational physics and include simulations of material properties using Monte Carlo and molecular dynamics techniques, and more recently, calculations of ocean wave properties. He completed a year's sabbatical last year at the Rosenstiel School of Oceanography in Miami, Florida. He is also interested in applications of computer technology in teaching. He is the co-author of a textbook for advanced classical mechanics (R. S. Jones and B. Hawkins, *Classical Mechanics Simulations*, John Wiley Sons, 1995) that is based on a set of computer simulations developed by the authors. The book has been translated into Japanese. He is also a senior investigator on an NSF-sponsored program at Loyola (the Internet Science Institute) to develop education materials for the World Wide Web.

Professor Raul Miranda

University of Louisville
Presently on leave as program manager for the catalysis program at Basic Energy Sciences, Chemistry Division, Office of Science, U.S. Department of Energy
Address: Chemical Engineering, Louisville KY 40222
Email: rmiranda@louisville.edu
URL: http://homer.louisville.edu/~r0mira01/resume.htm

At the start of this study, Raul Miranda was Director of the Kinetics, Catalysis, and Molecular Processes Program within the Chemical and Transport Systems Division of the NSF's Directorate for Engineering. He was on leave from the Chemical Engineering Department of the University of Louisville, where he has been a faculty member since 1983 and to which he returned in 2000. His industrial experience includes work at Compania Minera Aguilar, Argonne National Laboratory, and Pittsburgh Energy Technology Center.

Dr. Miranda's research interests focus on heterogeneous catalysis and solid-state materials. In the field of catalysis, he has conducted research on hydrotreatment and selective oxidation using transition metal oxides, chalcogenides, and solid acids. His materials research has involved sol-gel and solid-state synthesis as well as microfabrication of chemical devices and chemical sensors. Research in molecular modeling includes *ab initio* models of solid surfaces and surface reactions, and molecular dynamics of gas diffusion. Finally, his research involving computer data analysis has examined data mining in chemical process databases and neural-network property predictors.

Dr. Anna Tsao (European site visits)

Center for Computing Sciences, Institute for Defense Analyses
Address: 17100 Science Drive, Bowie, MD 20721
Email: anna@super.org

Anna Tsao received her BA (1972) and MS (1974) in mathematics from Michigan State University and her PhD degree in mathematics from the University of Michigan (1981). She has held academic positions at the U.S. Naval Academy and Texas Tech University, and industrial positions at Hughes Aircraft Company, AT&T Bell Labs, and the Institute for Defense Analyses. Since 1987, her research has focused on scalable, architecture-aware numerical linear algebra algorithm design, particularly for large-scale eigenvalue problems. She has published numerous papers in mathematics and computer science, organized or co-organized several mini-symposia and workshops, and is currently a member of the editorial board for the Frontiers in Applied Mathematics book series.

From 1994 to 1998, she was a program manager for Applied and Computational Mathematics at the Defense Advanced Research Projects Agency (DARPA). While at DARPA, she conceived and obtained DARPA funding for four major multidisciplinary thrusts: Physically Based Signal/Image Processing, Virtual Integrated Prototyping (VIP) for Vapor Deposition of Functional Thin Films, Optimized Portable Application Libraries (OPAL), and Virtual Electromagnetic Test Range (VET). Mathematical research she funded, managed, and championed resulted in broad technological and industrial impact inside and outside the Department of Defense. Areas in which her programs were recognized include data compression, automatic target recognition, computational electromagnetics, computer science, and microelectronics and materials processing.

Her current professional activities include consulting for DARPA and the NIH National Institute on Alcohol Abuse and Alcoholism.

APPENDIX B.1. SITE REPORTS—EUROPE

Site: **Air Liquide**
1 Chemin de la Porte des Loges
Les Loges-en-Josas, France
Tel: +33-1-3907-6363; Fax: 33-1-3907-6104
http://www.airliquide.com/

Date Visited: 15 September 1999

WTEC Attendees: P.T. Cummings (report author), A.M. Chaka, K.K. Irikura, P. Vashishta, P.R. Westmoreland

Hosts: Dr. Marc Till, Domain Manager, Modeling (marc.till@airliquide.com)
Dr. Pluton Pullumbi, Molecular Modeling (Group Leader, pluton.pullumbi@airliquide.com)
Ms. Sandrine Fudaley, Molecular Modeling (PhD student)
Ms. Stephanie Girard, Molecular Modeling (PhD student)
Dr. Stephane Mélen, Computational Fluid Mechanics

BACKGROUND

Air Liquide is the world's largest industrial gases company, with sales of 39.9 billion French francs (FF) in 1998. Research and development activities were FF 874 million in 1998. Air Liquide operates in over 60 countries and has 27,600 employees. Its current share of the industrial gases market is just over 20%; this will increase to over 30% with the acquisition of half the assets of BOC, which has been bought in roughly equal parts by Air Liquide and U.S.-based Air Products and Chemicals.

MOLECULAR MODELING R&D ACTIVITIES

Air Liquide's research and development spending totaled FF 874 million in 1998. It has seven research centers (St. Ouen L'Aumons, near Paris, established 1958; Les Loges-en-Josas, established 1970; Chicago, established 1983; Japan, established 1986; Castres, France, established 1986; Newport, Delaware, United States, established 1988; Hamburg, established 1996). The research center at Les Loges-en-Josas functions as a central research laboratory. In 1998, Air Liquide filed 1,100 patents to protect 212 inventions. This is representative of the annual patent and invention activity at Air Liquide.

Air Liquide's research is mostly aimed at applications of industrial gases, to expand existing markets and create new markets. Its research is broken into 12 domains:

1. process chemistry
2. metals and heat treatment
3. combustion
4. bioresources
5. analysis
6. welding and cutting
7. medical
8. electronics
9. plasma
10. modeling
11. process control
12. gas production

The last three are cross-cutting, while the first nine reflect business units within Air Liquide. The modeling domain includes computational fluid mechanics and process design, as well as molecular modeling and simulation (MMS). The modeling domain has about 100 active university collaborators throughout the world.

MMS began as a five-year feasibility experiment at Air Liquide in 1994. The evaluation period is drawing to a close and MMS will now become an ongoing modeling activity at Air Liquide. It is expected that MMS will reach the level of maturity and acceptance that computational fluid mechanics (CFD) has already achieved at Air Liquide. The CFD group was established 20 years ago and now has six people.

The MMS group currently consists of three people—the group leader, Dr. Pullumbi, and two junior members (Ms. Fudaley and Ms. Girard, who are both pursuing PhD degrees in molecular modeling while working full-time at Air Liquide). Research projects in MMS can be divided into three categories: short-term, in response to a customer need or internal Air Liquide needs; mid-term, funded by Air Liquide internal clients; and long-term, funded centrally. Most of the MMS projects are long-term.

In comparison with its major competitors, Dr. Pullumbi's perception is that Air Products does more molecular modeling, a consequence of it also having its chemicals business. Praxair has three to four people, about the same size group as Air Liquide's. BOC has had no molecular modeling activity.

It was stated that reduction of the results of Air Liquide MMS activities to a design rule and/or a quantitative structure-activity relationship (QSAR) is most desirable, since these can be utilized most effectively by others in the company.

One of the main activities of the MMS group at Air Liquide is to design new adsorbents. The typical modeling sequence for the design of a new adsorbent at Air Liquide is:

1. Develop understanding of adsorption at the molecular level: This includes using molecular mechanics methods for structure refinement and characterization, hybrid quantum mechanics/molecular mechanics methods for predicting topology and stability, and Monte Carlo methods for predicting adsorption properties.
2. Identify key molecular features and propose a consistent theory: This involves developing a QSAR to extract information from experimental data, developing appropriate descriptors for estimating adsorption properties, and comparing to experiment to validate the correlation.
3. Molecular modeling to design new materials with improved properties: This consists of generating new materials using combinatorial chemistry algorithms, using QSAR to predict adsorption properties of new materials, and finally verifying predictions experimentally.

Other examples of MMS activities at Air Liquide include ceramic-membrane separations (for manufacturing synthesis gas from natural gas) and temperature-swing adsorption. In the latter case, the goal is long-term effectiveness for oxygen adsorption at elevated temperatures and at high partial pressures of water. Binding energies were calculated for oxygen with cobalt, iron, manganese, and copper organometallics. By varying ligands, changes in the binding strength were analyzed.

Since computational quantum chemistry (CQC) is used extensively at Air Liquide, the need for Order(N) methods was emphasized as the most outstanding technical issue.

CONCLUSIONS

MMS activities at Air Liquide are making the transition from being a trial effort to becoming an established component of Air Liquide's modeling activities. Given the small size of the group and the relatively short period of its existence, the group has made an impact in a remarkably broad range of areas. The small size of the group means that it necessarily relies heavily on commercial MMS software, rather than developing in-house codes. The impression gained is that MMS activities will increase in Air Liquide as the value of the methods is more widely appreciated within the company. It appears that MMS today has the status that CFD

once had, and just as the latter has grown in size and importance over the past two decades, the MMS activities can be expected to do the same.

Table B1.1
Technical Summary of Molecular and Materials Modeling at Air Liquide

Past and Present Applications	Important Emerging Applications	Theoretical Approaches & Software
Adsorptive separations Permeability separations VLE and phase transitions Adsorptive materials Separation polymers Novel ligands Catalysis Process modeling Chemical conversions New ideas Complexation of aqueous lanthanide ions (Pullumbi et al. 1997)	Mesoscale modeling	Molecular graphics, molecular mechanics/molecular dynamics Monte Carlo Quantum mechanics Electronic density Functional theory QSAR Combinatorial chemistry Data management Mainly MSI software

REFERENCES

Pullumbi, P., and J. Ligniéres. 1998. Molecular modeling of gas separation in zeolites. *Proceedings of the AIChE Topical Conference on Applying Molecular Modeling and Computational Chemistry*, Miami Beach, November 15-20, paper 76b.

Pullumbi, P., R. Guilard, C. Mellot, and J. Ligniéres. 1997. Prediction of the coordination scheme in Ce-tetraaza tetra propionic acid complexes using the molecular electrostatic potential to study the initial stages of complexation. *Solutions*, MSI, Nov. 1997; http://www.msi.com/solutions/cases/notes/liquide_full.html.

Pullumbi, P. 1998. Molecular modeling of transport properties of polymers for gas separation membranes. *Proceedings of the AIChE Topical Conference on Applying Molecular Modeling and Computational Chemistry*, Miami Beach, November 15-20. American Institute of Chemical Engineers, paper 52d.

Site: **BASF AG**
Zentralbereich Informatik
ZOI/ZC-C13; 67056 Ludwigshafen
Germany
http://www.basf.com/

Date Visited: 15 September 1999

WTEC Attendees: R. Miranda (report author), E. Stechel, A. Tsao, P. Kollman, K. Morokuma

Hosts: Dr. Ansgar Schäfer, Chemist, Central Informatics; ansgar.schaefer@basf-ag.de; Tel. 49-621-60-78376; Fax 49-621-60-49463
Dr. Gerhard Lippert, Chemist, Central Informatics
Dr. Gerhard Cox, Physicist, Ammonia Laboratory, Analytical Division
Dr. Jens Sadowski, Chemist, Main Laboratory
Dr. Wolfram Alenhaufen, Chemist, Main Laboratory, Pharmacy Division
Dr. Eric Hädicke, Chemist, Polymer Laboratory

BACKGROUND

BASF is a leading chemical manufacturing group with sales in excess of DM 54 billion and a portfolio of about 8,000 products. It has 106,000 employees worldwide—43,000 at Ludwigshafen—and owns 400 companies. Its products range from oil and natural gas, to basic, intermediate and specialty chemicals, colorants and coatings, plastics and fibers, crop protection products and pharmaceuticals.

Badische Anilin & Soda Fabrik was founded in 1865 in Ludwigshafen to produce coal tar dyes. In 1901 the company initiated fertilizer production and in 1913 commercialized the first Haber-Bosch ammonia synthesis process. In 1929 it launched the industrial synthesis of styrene and the beginning of the plastics era. In 1963 BASF industrialized new synthesis processes for vitamin A and the herbicide Pyramin. Most recently, in 1998, BASF commercialized Sibutramine in medicines for pathological obesity.

Since 1965, BASF has expanded into about 400 companies around the world. The main sites are in Ludwigshafen (Germany), Antwerp (Belgium), Tarragona (Spain), Freeport and Geismar (United States).

Its largest segments in terms of sales are plastics and fibers (27%) and colorants and finishings (22%), chemicals (19%), health and nutrition (18%), and oil and gas (10%). The fastest growing segment is health and nutrition, increasing 11% per year in terms of sales. Major activities in these segments include the following:

- Health and nutrition—Pharmaceuticals: drugs for treating obesity, diseases of the cardiovascular system, thyroid insufficiency, and disorders of the central nervous system. Fine chemicals: vitamins, carotenoids, amino acids, enzymes, acids, flavors and fragrances, polymers and fine chemicals for the pharmaceutical, food and cosmetics industries, and animal nutrition products. Crop protection: herbicides, fungicides, and growth regulators.
- Colorants and finishing products—colorants: dyes for textiles, pigments for plastics, printing inks and plates. Dispersions: acrylic acid and acrylates, acrylate and styrene/butadiene dispersions, paints, coatings and adhesives, chemicals for the paper industry, and superabsorbents. Coatings: automotive finishes, industrial and building paints.
- Chemicals—petrochemicals and inorganics: ethylene, propylene, benzene; basic chemicals such as methanol, acetylene, ammonia, sulfuric acid; and catalysts. Fertilizers: straight nitrogen fertilizers and compound fertilizers; products for plant care. Industrial chemicals: plasticizers such as 2-ethylhexanol, phthalic anhydride and higher oxo alcohols, oxygenated solvents such as butanol; adhesive resins, formaldehyde and melamine. Intermediates: amines, diols, carboxyl and dye intermediates; carboxylic acids. Specialty chemicals: alkylene oxides, glycols, surfactants, detergents, fuel and lubricant additives, biocides, electroplating chemicals.

- Plastics and fibers—styrenic polymers: polystyrene, expandable polystyrene and specialty foams. Engineering plastics: copolymers such as ABS and SAN, nylons, PBT, POM, and thermoplastics.
- Polyurethanes: polyols and isocyanates, polyurethane elastomers, PVC. Fibers: caprolactam, adipic acid, hexamethylenediamine, nylon fibers. Polyolefins: polyethylene (with Hoechst) and polypropylene (with Shell).
- Oil and gas—exploration, production, refining and transportation. Run by Wintershall AG.

R&D ACTIVITIES

The BASF central research complex is located in Ludwigshafen, which also serves as the coordination center for research worldwide. The Research Executive Director for global research is Dr. Stefan Marcinowski. BASF has decentralized research and development departments primarily at sites in Germany, Great Britain, United States, and Japan.

The Ludwigshafen center covers the following:

- Chemicals, catalysts, and process engineering
- Polymers
- Active ingredients for pharmaceuticals and crop protection
- Specialty and performance chemicals

Pharmaceutical research is carried out mainly at Knoll AG and subsidiaries worldwide. Crop protection products are developed at the Limburgerhof Agricultural Center and other stations around the world. Polymer development is carried out at various sites; coatings and crop protection mainly in Germany and the United States; and fibers in the United States. BASF employs about 10,000 research personnel, of whom 8,000 are located in Germany. Health and nutrition receives 49% of BASF's total research investment of DM 2.8 billion, followed by plastics (14%), colorants (12%) and chemicals (12%). The other sectors receive less than 5% each. The outcome of BASF's research activity is represented by its ownership of over 90,000 patents.

BASF strategy for the next few years is to increase their commitment to biotechnology and pharmaceuticals and maintain innovation in the other segments, making strategic use of joint ventures. Increased attention to plant genomics will result in new ways to increase crop yields. For example, a new company called BASF Plant Science GmbH will combine seed development and plant biotechnology.

In another new company, BASF-Lynx Bioscience AG, microorganisms are genetically optimized for fermentation processes. In pharmaceuticals BASF and Hoechst are developing an active substance for the treatment of cardiovascular disease, and together with Novartis, these companies are developing a compound for the treatment of schizophrenia. In the chemical segment, BASF, DaimlerChrysler and Ford have developed a catalyst for hydrogen production from methanol for use in automobile fuel cells. With GE Plastics, BASF will develop lightweight automobile bodies. Together with detergent companies, BASF is producing new environmentally safe surfactants for cleaning products and better fuel additives.

The Ludwigshafen center houses the Ammonia Laboratory, which carries out most chemical, catalysis, and process engineering research. Some of its recent developments have been caprolactam from butadiene, acrylic acid, vinyl ethers and hexandiol. It employs about 1,800 persons, of whom 350 are scientists and engineers. Ludwigshafen also houses the Main Laboratory, which is focused on life sciences research. Currently it targets new drugs for cardiovascular diseases, new fungicides, and food additives. It identifies molecular targets and develops high throughput screening as well as toxicity tests. The Main Laboratory employs 1,300 persons, including 200 scientists and engineers. There is cooperation with academic groups worldwide. Other laboratories in Ludwigshafen include the Polymers Laboratory, the Colorants Laboratory, and the Engineering Research and Development Laboratory. BASF's total number of researchers with PhDs in science or engineering is about 2,500, of whom 1,000 are in Ludwigshafen.

Of these, the number of molecular or materials modelers is 12-15, five of whom are in the Central Informatics Department and the rest are distributed among the research laboratories. The Central Informatics Department in addition has 10 researchers who work on statistics, fluid dynamics, optimization, and reactor design.

STAFF EXPERTISE

Dr. Schäfer

As part of the Central Informatics Department, Dr. Schäfer has carried out molecular quantum chemistry calculations for a wide variety of applications. About 60% of them have been DFT calculations for homogeneous catalysis. He also uses computational chemistry for thermochemistry and spectroscopy and hybrid QM-MM for large-scale homogeneous, heterogeneous and enzymatic catalysis.

Dr. Lippert

Dr. Lippert recently graduated from the Parrinello group and now is employed in the Central Informatics Department. He carries out DFT and CP-MD calculations for catalysis, mainly heterogeneous.

Dr. Cox

Dr. Cox carries out high-resolution microscopy and X-ray diffraction, and uses direct methods for structure calculation and refinement. His main applications are bulk materials and heterogeneous catalysts for basic and intermediate chemical products. The analytical division of the Ammonia Laboratory employs one PhD full-time modeler.

Dr. Sadowski

As part of a group of three modelers and two crystallographers, Dr. Sadowski's main responsibility is 3-D structure prediction of proteins for medicinal chemistry and crop protection. He also dedicates 30% of his time to method development for *de novo* design and combinatorial library design.

Dr. Altenhaufen

Dr. Altenhaufen has expertise in biophysics, ion channels, electrophysiology, and ligand-protein interaction, for applications in life science and agriculture. He uses MM for structure prediction, design, and docking calculations.

Dr. Hädicke

Dr. Hädicke worked for the last nine years with the Polymer Laboratory, which has also two other PhD-level modelers dedicated to polymer structure and property calculations. He collaborates closely with an experimental chemist. They use MM, MC, MD, and Mesodyne for large-scale systems. Their applications are in colorants and pigments, latex particles, additives, vitamins and nutritional products.

They also use DFT, CC and ZINDO methods for catalysis and spectroscopy. They have developed QSPR for solubility parameters of polymers.

METHODOLOGY FOR MOLECULAR MODELING

BASF's researchers have a range of molecular modeling expertise, from classic to *ab initio* quantum chemistry, from atomistic to mesoscale and continuous, from fluids to solids. They have particular strengths in thermochemistry, spectroscopy, protein structure prediction, protein-ligand interaction, homogeneous catalysis, polymer rheology and structure-property prediction, and a limited amount of heterogeneous catalysis. Historically, they started in life-science-related modeling around 1985, then grew into polymers and homogeneous catalysis starting around 1990 (http://www.msi.com/solutions/cases/notes/reactions.html)

and into meso and large-scale quantum chemistry starting around 1995; heterogeneous catalysis modeling is in the early stages.

For thermochemistry and spectroscopy, they rely on DFT and Coupled Cluster theory methods, while for catalysis they employ hybrid MM-QM and Car-Parrinello MD. For protein structural applications, they use a 3-D structure generator with validation provided by a well-established X-ray crystallography group. In terms of bulk material including heterogeneous catalysts they use high-resolution microscopy and X-ray diffraction. For polymers, structural data is obtained with X-ray crystallography, and property prediction is done with MM-MD, MC and QSPR. In particular, they use Mesodyne for mesoscale properties of aggregates.

Thus they are able to model inhibition properties by interaction of polymeric additives with surfaces, solubility parameters, mixing compatibility, stress-strain relations (http://www.msi.com/solutions/cases/notes/micelles.html, Surfactants). They also model pigments with MM and Mesodyne and color absorption properties with ZINDO (http://www.msi.com/solutions/cases/notes/perylenered.html, Crystal habit in pigments).

SOFTWARE

BASF is member of all MSI consortia. BASF researchers find MSI software invaluable for visualization. The new formulation consortium will provide added value particularly for its data-mining and QSPR prospects. Since BASF maintains several active collaborations with universities, they prefer academic codes for 90% of the numerically intensive computations.

HARDWARE

BASF molecular modeling applications are run on central and distributed computing resources. Central Informatics maintains an SGI 16-processor Power Challenge, a 32-processor Origin and an IBM 10-processor SP2. For graphics they rely on SGI octanes. The research laboratories also maintain a SGI 12-processor Power Challenge, an IBM 32-processor SP3, and Linux clusters of 8 double processors. The Polymers Laboratory also runs a 100-NT Linux cluster overnight.

OPERATIONAL AND STRATEGIC ISSUES

Eighty percent of molecular modeling projects at BASF are coordinated and sometimes financed by Central Informatics. Those projects are typically for service to internal customers in the research laboratories, who pay for the calculation and interpretation. The other 20 percent usually involve method development. Those three-year research projects are subject to review and compete internally for funding. In the case of service projects, the modeler is approached by the customer and ideally is involved in the research from the beginning as part of the team. In order to maintain its reputation as an essential resource for the company, the modeling team is forced to maintain a very high rate of success. Therefore, it must cautiously select the projects to work on according to (a) possibility of success, (b) resource availability, (c) state of development of the project.

BASF management usually measures success in terms of the savings achieved in process or product development. For the modeler, success is usually quantified by the number of patents co-authored. Each modeler is involved in 2-5 major projects. Collaborations with academia are common, so PhD students or post-docs financed by BASF are also included in modeling projects, particularly those of method development.

NEEDS FOR FUTURE DEVELOPMENTS

The BASF hosts identified several current needs and future opportunities for molecular modeling. In quantum chemistry, DFT will remain the main method for the next five years and they foresee higher level

ab initio techniques and linear scaling for cluster calculations. Catalysis applications would benefit from more reliable functionals and semi-empirical approaches for validation. They cited collaborations with Wemer (Molpro) and Ahlrich (Turbomole) for improving high-level multi-reference methods.

For heterogeneous catalysis the primary needs are to identify mechanisms of reaction and to improve catalyst performance. Thus, large-scale Order(N) schemes are in demand. Calculations should be automatically set up by means of template interfaces and run combinatorially in parallel. Since atomically correct structural information on heterogeneous catalysts is scarce, zeolites, which are better known than other catalysts, should be used whenever possible. In terms of most bulk structures, new and reliable methodologies are needed to solve structures from powder diffraction data and to interpret images from electron microscopy. Simulated annealing, a direct method, should be preferred over refinement.

In the area of protein design, empirical (COMFA) calculations are carried out for docking and scoring. For drug design, they would like to see small academic applications, although they make use of MSI, Tripos, FlexX, and LUDI (developed at BASF). Improved scoring functions, and in the long term, MM/QM approaches for exact binding energy, are expected. In a joint venture with the Lynx company, BASF is improving those functions using the wealth of genomics and protein X-ray data.

Researchers in the pharmaceuticals group needed a Web-based approach to create 3-D data exchange. They developed the CrunchServer to be used by bench chemists to calculate thermodynamics of organic reactions, structure optimization by DFT, and energies by MP2 (http://www.scsc.ethz.ch/~flukiger/Speedup/, http://pubs.acs.org/hotartcl/cenear/971006/ab.html). However, it remained unused because of inefficient data management. Currently 3-D data are shared using WebLab, freeware from MSI, and Chime, a Web browser plugin.

For polymers simulation, the usual limitation is time and length scale. They need relaxation times of up to a second, as well as mesoscopic systems for molecular structure-property calculation: folding, stress-strain, shear flow, morphology after extrusion. They also need to include realistic force fields and polyelectrolytes. Collaboration with the MPI at Mainz is focused on the prediction of polymer morphology and properties from monomer structure and force fields.

For fast screening of proteins or catalysts, there is need for crude filters, virtual library design, and data-mining. BASF has spun off a combinatorial chemistry company called HTS that is developing those approaches.

CONCLUSIONS

The use of molecular modeling and simulation (MM&S) at BASF is relatively recent when compared with the long history of the company and its many successes, most of which have been the result of product and process development in the traditional laboratory environment. The current mode of operation, which is mostly based on service or problem solving projects, lends itself to slow growth. The necessity to constrain the type of problems to obviously tractable ones in order to avoid failures prevents exposure to more groundbreaking approaches, and hence reduces the impact of MM&S on the company output. Certainly the increasing relevance of combinatorial chemistry to drug design will benefit MM&S because of its clear role of helping library design by elimination of members computationally shown to be inactive. The clear success of the polymers modeling group is due to a repeatedly demonstrated ability to predict properties on the basis of molecular structure and processing conditions. The group includes close cooperation with experimentalists, which ensures higher availability of relevant data. Joint publication with experimentalists is perceived as valuable.

In terms of strategy, the modeling groups encourage the generation of reliable and high-resolution structural information as well as make use of available genomics information. They also encourage the dissemination of 3-D structural results and the input of 3-D data directly from the users. They also interact with external colleagues. The long-term development projects are run in collaboration with academic groups, while the BASF modelers concentrate on identifying and solving fruitful applications.

Site: **Bayer AG**
http://www.bayer.com/

Date: 4 February 2000, 30 November 2000

WTEC reporter: P.R. Westmoreland (report compiled from literature sources except for SC segment)

INTRODUCTION

Bayer AG (http://www.bayer.com/) is an international chemical company with a wide spectrum of products. In order of sales volume, which totaled $25 billion in 1998, its principal business segments are as follows (Bayer n.d.):

- Polymers (32%)—Thermoplastics, rubber, polyurethanes, fibers, cellulosics, coatings, adhesives, and pigments
- Healthcare (25%)—Prescription and generic drugs, medical diagnostics
- Chemicals (16%)—Fine chemicals, basic chemicals, petrochemicals, specialty chemicals, polyethylene, metal and ceramic powders, solar-cell-grade silicon, fragrances and flavors
- Agriculture products (11%)—Herbicides, pesticides, animal healthcare

Bayer's main manufacturing operations are in Germany, notably at its headquarters site in Leverkusen (25,000 people). Its Uerdingen site manufactures fine chemicals, specialty chemicals, polycarbonates, and inorganic pigments (the world's largest supplier). Around the world, the company's manufacturing, research, and marketing operations employ 120,000 people.

APPLICATIONS

Bayer has carried out a wide range of applied molecular and materials modeling, including drug development, crystal habit for organics (MSI 1995), modeling solvation effects in aqueous solutions (Andzelm et al. 1995), atmospheric kinetics (Frank et al. 1998), homogeneous catalysis (Lohrenz 1998), UV/Vis spectra (Klampt 1996), inclusion in process simulators (Brüll 1996), and life sciences modeling (Mierke et al. 1995; Cremer et al. 1994; Müller et al. 1995; Böhner et al. 1996; Kessler et al. 1996; Lutz et al. 1997; Müller et al. 1997; Müller and Giera 1998; Gurrath et al. 1998; Cramer et al. 1999; Müller n.d.). Important modeling developments at Bayer have included the COSMO and COSMO-RS methods marketed by MSI.

One published example was work to reduce moisture hold-up in an aromatic disulphonic acid being sold by the company (MSI 1995). Because the crystal habit was flat and plate-like, a more three-dimensional, isometric crystal was sought that would be easier to dry. The work proceeded in two stages, both involving a mixture of computation and experiment. First, growth of the natural crystal was analyzed using software from Biosym (now MSI) for crystal morphology modeling and for analysis of diffraction data. Its plate-like shape was quantitatively interpreted to be the result of slow growth on one of the crystal faces. Second, additives were suggested that would bind more favorably (energetically) to the fast-growing faces, slowing growth, or to the slow-growing face, speeding growth there. An aromatic trisulphonic acid appeared especially favorable from computation. Subsequent experiments bore out the prediction, and moisture hold-up in the improved crystals was reduced by a factor of four.

Other published applications include catalysis and use of results in process design. Dr. John Lohrenz (a PhD chemist originally from Tom Ziegler's group in Canada) reported the use of DMol and COSMO-RS to study the uncatalyzed and ZnCl2-catalyzed SN2 reaction of Cl- with methanol in water. Heterogeneous catalysis and adsorption for O2 on silver have also been studied in the contexts of selective ethylene partial oxidation and of oxydehydrogenation of methanol (Lohrenz 1998).

A site in Bushy Park, South Carolina (United States) manufactures pigments. The focus of molecular modeling at that site is pigment crystal binding energy. Dr. Gamil Guirgis (Gamil.Guirgis.B@bayer.com) began this work in mid-1999 using programs of Prof. Allan S. Myerson (Polytech Univ. of Brooklyn). Current calculations are done using MSI's Cerius2 programs. The main limitations for their calculations determining the single-crystal structural parameters and identifying the appropriate force field to use in building the morphology. *Ab initio* molecular-orbital methods have proven very useful to calculate the charge distribution for more accurate results.

In October 2000 (http://www.tripos.com/about/press/2000/2000.10.16.html), Lion Biosciences AG and Tripos, Inc. announced that Bayer was investing $25 million to construct an integrated cheminformatics technology platform, intended to speed Bayer's identification of lead candidates for its drug and agricultural chemical programs.

REFERENCES

Andzelm, J., C. Kölmel, and A. Klamt. 1995. Incorporation of solvation effects into density functional calculations of molecular energies and geometries. *J. Chem. Phys.* 103, 9312-9320.

Bayer AG. n.d. Company Web site: (http://www.bayer.com/).

Brüll, L. 1996. Optimization of entire chemical production plants, computational chemistry and chemical engineering. *Proceedings of the Third UNAM-Cray Supercomputing Conference*, World Scientific, pp. 367-376; http://www.kfa-juelich.de/compchem/proceedings/Bachmann/bachmann.html.

Cramer, J., S. Fiori, G. Müller, C. Renner, S. Pegoraro, and L. Moroder. 1999. Comparison of a Monte Carlo strategy with a combined DG/MDSA method for structure determination of bicyclic peptides. *J. Mol. Model.* 5, 287-295.

Cremer, D., L. Olsson, F. Reichel, and E. Kraka. 1994. Calculation of NMR chemical shifts—the third dimension of quantum chemistry. *Isr. J. Chem.* 33(4), 369-85.

Frank, I., M. Parrinello, and A. Klamt. 1998. First-principles simulation of chemical reactions: the mechanism of the gas-phase reaction of OH radicals with ketones. *J. Phys. Chem.* A 102, 3614.

Gurrath, M., G. Müller, and H.-D. Höltje. 1998. Pseudoreceptor modelling in drug design: Applications of Yak and PrGen. *Perspectives in Drug Discovery and Design* 12/13/14, 135-157.

Kessler, H., R. Gratias, M. Hessler, M. Gurrath, and G. Müller. 1996. Conformation of cyclic peptides—Principle concepts and the design of selectivity and superactivity in bioactive sequences by "spatial screening." *Pure and Applied Chemistry* 68, 1201-1205.

Klamt, A. 1996. Calculation of UV/Vis-spectra in solution. *J. Phys. Chem.* 100, 3349.

Lohrenz, J. 1998. Quantenchemie in der chemischen industrie: Von der berechnung von molekülstrukturen bis zur simulation von chemischen reaktionen. Höchstleistungsrechnen in der Chemie, Forschungszentrum Jülich, February 15-18; http://www.zam.kfa-juelich.de/compchem/proceedings/Lohrenz/lohrenz.html.

Lutz, J., R. Romano-Götsch, C. Escrieut, D. Fourmy, B. Matthä, G. Müller, H. Kessler, and L. Moroder. 1997. Mapping of the ligand binding site of cholecystokinin-b/gastrin receptor using lipo-gastrin peptides and molecular modelling. *Biopolymers*. 799-817.

Mierke, D.F., S. Rudolph-Böhner, G. Müller, and L. Moroder. 1995. Structure of two microcystins: Refinement with NOEs and ensemble calculations. *Biopolymers* 36, 811-828.

MSI. 1995. http://www.msi.com/solutions/cases/notes/bayer.html; abstracted from A. Karbach, *mc2 Magazine.* 3(1), 1.

Müller, G. N.D. Towards 3D structures of G protein-coupled receptors: A multidisciplinary approach. *Curr. Med. Chem.* (invited review in press).

Müller, G., A. Feriani, A.-M. Capelli, and G. Tedesco. 1995. The impact of modern multidimensional NMR-spectroscopy on rational drug design. *Chimica Industria* 77, 937-957.

Müller, G., and H. Giera. 1998. Protein secondary structure templates derived from bioactive natural products: Combinatorial chemistry meets structure-based design. *J. Comp.-Aided Mol. Design* 12, 1-6.

Müller, G., M. Gurrath, and H.-D. Höltje. 1997. Molecular dynamics simulations on G protein-coupled receptors—I: Implementation and validation of a membrane mimetic. *Proceedings of the 2nd International Conference on Molecular Structural Biology* (A.J. Kungl, P.J. Andrews, Eds.), Wien, p. 29.

Gurrath, M., G. Müller, H.-D. Höltje, and L. Moroder. 1997. Molecular dynamics simulations on G protein-coupled receptors—II: Application on bacteriorhodopsin and CCK-B receptor complexes. *Proceedings of the 2nd International Conference on Molecular Structural Biology* (A.J. Kungl, P.J. Andrews, Eds.), Wien, p. 124.

Rudolph-Böhner, D.F.S., G. Müller, and L. Moroder. 1996. Structure refinement of microcystin with ensemble calculations and NMR data. In *Peptides: Chemistry, Structure and Biology* (Kaumaya, P.T.P., and R.S. Hodges, eds.), Mayflower Scientific Ltd., Kingswinford, pp. 724-726.

Coverage:	**Belgium: University of Antwerp, Janssen Pharmaceutica, Solvay**
Visit location:	Workshop at Eindhoven University of Technology
Date:	17 September 1999
WTEC Attendees:	A.M. Chaka (report author), K.K. Irikura, P. Vashishta, P.R. Westmoreland
Participants:	Dr. Anik Peeters, Senior Research Assistant, Structural Chemistry Group, Chemistry Department, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium, 32-3-8202366, fax 32-3-8202310, anik@uia.ac.be Prof. Rutger Van Santen, Eindhoven University of Technology, R.A.v.Santen@tue.nl

BACKGROUND

This report on Belgian activities is organized somewhat differently than the other site reports because it is derived from the literature, the Web, and discussions at workshop at the Eindhoven University of Technology (*Technische Universiteit Eindhoven*, TUE).

MODELING ACTIVITIES AT THE UNIVERSITY OF ANTWERP

The University of Antwerp is a relatively young university, founded in 1971 from three former universities: RUCA, UFSIA and UIA. There are two main directions in the Structural Chemistry group (http://sch-www.uia.ac.be/struct/), synthesis and theory, led by Professors H.J. Geise and C. van Alsenoy, respectively.

The specific research interests of Prof. Geise are:

- Synthesis of oligomers
- Use of oligomers in organic light emitting diodes (LEDs) and electronic noses
- Applied quantum chemistry (as a tool for problem-solving in synthesis)

He began working in polymers but found that oligomers performed equally well or better for the applications of interest. He had started his career and research program with purely experimental chemical synthesis but eventually moved to the use of theory in 1994 to interpret spectra and explain results. Now all experimental students in the group are taught how to use Gaussian, what methods to use when, and the limitations of each method. They feel strongly that it is good to teach non-experts. Next week, a PhD degree will be granted to a student who did a combination of theory and experiment. The student, Frank Blockhuys, started his PhD work doing only experimentation, but later did some *ab initio* calculations to answer some of the questions that arose from his experiments.

Professor Geise is very interested in applications and is now involved in a company selling electronic noses. His use of computational chemistry is very much driven by the need to answer questions. Prediction methods that they see particular needs for at present are solubility/miscibility of organic compounds and polymers, adsorption on surfaces, transfer properties of charge carriers from electrodes, and fluorescent properties.

One example of a successful application of theory was the use of the molecular electrostatic potential of a thiophene derivative that led to better experimental conditions in the synthesis of a new poly(thienylene-vinylene) compound with a higher yield (Blockhuys et al. 1998). They employ the range of computational quantum chemistry from semi-empirical to electronic DFT and wavefunction-based *ab initio* calculations, as well as QSAR correlations.

Technically, Geise sees the most important modeling challenges for his group as being (1) the ability to do bigger problems, particularly to simulate solids, and (2) the ability to model transition metals and spectra more accurately. Theory must contribute, but the power of parallel computation is also very important.

Professor Geise feels that transfer of science to technology occurs most efficiently via a direct relationship between an industrial company and an academic research group. At the same time, government financial aid can be very useful to lower the barriers towards the (sometimes high) risks involved.

Professor van Alsenoy's research interests are in quantum chemistry, summarized as:

- *Ab initio* HF calculations of large systems (biomolecules, solid state)
- Modeling polymers and interactions with additives, prediction of elastic modulus
- Collaborations with Prof. V. van Doren (theoretical physics, RUCA), polymer engineering group (UCL) and industry

In Prof. van Alsenoy's group, they have written their own program, BRABO, based on Peter Pulay's program TEXAS, to perform Hartree-Fock (HF) calculations on large systems. A key development in their program is a fast method to evaluate 2-electron integrals, which van Alsenoy developed when he was a post-doc with Pulay. They have achieved linear scaling at the HF level, which enables them to do very large molecules. One of the largest systems they have done is Crambin, a protein that has 46 amino acids.

They do have MP2 energies but not forces in BRABO, so they use commercial codes (i.e., Gaussian) to supplement it. They feel it is not worth the time investment to code many of the post-Hartree-Fock techniques themselves, but they have implemented LDA DFT and are implementing other functionals as well (June 2001). Gaussian, however, has problems with geometry convergence for these large systems. There is also a bug in the redundant internal coordinates that has not been fixed. In BRABO, they monitor DIIS coefficients to achieve rapid convergence.

Another active area of research is the modeling of crystalline molecular materials using experimental X-ray structural parameters, then surrounding the molecule with point charges to investigate the differences between the gas and the solid-state environment.

The theoretical group has started a collaboration with an experimental polymer engineering group at the Catholic University of Louvain in Louvain-la-Neuve. In van Alsenoy's group they have analyzed the charges and dipole moments of additives, trying to predict which additive would bind more tightly to the polymer to give them a qualitative ranking. The recommended compound turned out to be too good a plasticizer. Statistical mechanics should be used to study the amorphous phase, and they are looking into ways of acquiring the expertise and programs, possibly through post-docs.

Industrial collaboration is also important to them. The theory group has a project with an unidentified industrial partner, who has funded the purchase of their computer. For this project, they have developed a strict timeline and list of deliverables. They applied for a grant from a government agency with the industrial partner but didn't get it because it was seen as too ambitious. They are reapplying and are confident about being funded because they now have enough results to convince the referees. (The project was subsequently approved.)

MODELING R&D AT JANSSEN PHARMACEUTICA N.V. (SUBSIDIARY OF JOHNSON & JOHNSON CO.) (Turnhoutsesteenweg 30; B-2340 Beerse, Belgium; http://www.janssenpharmaceutica.be/)

Janssen is one of the companies in Belgium with a strong modeling effort. The company actually has two groups, one of which does very focused project work; the other does more exploratory work, though it is still focused on deliverables. The exploratory group was initiated by Paul Janssen, who founded the company in 1953. He felt this type of research was essential, and the company is still doing research "in his spirit." He located the group in his former guest-house, where he would visit a few days a week, even after he retired. This group is led by Paul Lewi. They are working with Peter Kollman and often invite academics for an extended stay. The group has eight people doing modeling and several people doing experimental work as well. This is believed to be important to keep the research effort balanced. The group is doing well, and it is considered prestigious to have a position in this group. More information on the group can be found at the following Web sites: http://www.janssen-cilag.com/about/r&d/technology2.stm and http://www.janssen-cilag.com/about/janpharma/drpaul.stm. The other group (focused projects) is led by Jan Tollenaere.

Janssen Pharmaceutica was using Insight from Biosym/MSI but felt it was buggy, and new releases didn't fix the problems. Now they seem to be disenchanted with it and are considering developing their own code.

MODELING R&D AT SOLVAY (http://www.cam-com.com/solvay.htm)

Solvay Research is currently located in Brussels, having moved from Hanover, Germany. Professor van Santen had started collaborative work with Solvay when the company was still located in Hanover. In Hanover there was a small modeling group interested in developing a new heterogeneous catalyst for environmental reasons. They had hired a brilliant spectroscopist who also wanted to do modeling. A director believed in this person, and gave him strong support. They recognized that they would need external help to enable the spectroscopist to do advanced theory, and at that point, they requested Prof. van Santen to help. Prof. van Santen also spent a full day analyzing their experimental problem, which he feels is one of his key strengths. (He notes he could not do that effectively without the experience he gained at Shell.) Now Solvay has a new commercial catalyst that reduces chloride to chlorine. Prof. van Santen's name is included on the patent. He believes strongly that this catalyst would never have been developed without experiment working in parallel with theory.

When Solvay moved to Brussels, they continued to sponsor a theory program in van Santen's group for a period of time. Today, however, they sponsor more experimental work in selsisequiods that have the same properties as zeolites but are completely controlled on a molecular level. Prof. van Santen's group developed epoxidation catalysts in collaboration with them that have since been patented. In Hanover they were quite open and flexible, but the current culture in the Brussels organization is more bureaucratic.

Solvay currently has one person with a workstation who models Ziegler-Natta catalysts and metal clusters. He does quantum mechanics with MSI software.

OTHER MODELING R&D IN BELGIUM

A large number of companies in Belgium have most of their research located in Germany, like Bayer. Solvay is one of the few which has their major research facility in Belgium.

Professor Shoshona J. Wodak leads research on protein folding at the Université Libre de Bruxelles, Unité de Conformation des Macromolécules Biologiques (http://www.ucmb.ulb.ac.be/).

Dr. Luc Vanquickenborne of the Katholieke Universiteit Leuven leads quantum chemistry research on calculating gas-phase kinetics and solvent effects and on magnetic and optical effects (http://www.quantchem.kuleuven.ac.be/research/quantum/quantum_en.html).

Recently the research group of Prof. Paul Geerlings of the Brussels Free University (VUB) has started research in large biomolecular systems, in collaboration with an experimental group of the same university and the theoretical group of Van Alsenoy in Antwerp.

REFERENCE

F. Blockhuys et al. 1998. *J. Mol. Struct* 445, 187-195.

Site: **BG Technology Ltd. (now Advantica Technologies Ltd.)**
Ashby Road
Loughborough
Leicestershire LE11 3GR
United Kingdom
Tel +44 (0)1509 282116, Fax +44 (0)1509 283118
http://www.advanticatech.com/

Date: 30 December 1999

Report author: P.R. Westmoreland

Interviewee: Dr. David R. Bates, david.bates@advanticatech.com

BACKGROUND

With origins in British Gas, Advantica is a leading provider of technology and engineering services to customers in gas, pipelines and associated industries in over 30 countries worldwide. Advantica is part of the Lattice Group of companies, has a turnover of £80 million and employs 800 people in the UK. The company is currently pushing into new markets, using the skills developed for the complex gas industry in Great Britain to solve problems and deliver solutions in others.

R&D ACTIVITIES

Advantica (formerly BG Technology) first used molecular modeling to study gas separation by zeolites. This work began with sponsored university research in the late 1980s, followed by in-house studies from the mid-1990s. Molecular simulations have been emphasized, analyzed mostly with MSI Cerius2 software on SGI workstations. At present, one in-house specialist (Dr. Bates) carries out this type of work.

An early success was in simulation of X-ray diffraction data for core samples of rock (Bates 1995). Fractional mineral composition was deduced from core-sample data by Rietveld refinement of combinations of single-phase X-ray diffraction patterns.

Gas separations continue to provide most of the topics of study. In collaboration with Birkbeck College, University of London, ion exchange and dehydration of a zeolite (clinoptilolite) were simulated using a combination of GCMC and MD techniques (O'Connor et al. 1998). This study revealed a possible explanation of hitherto anomalous gas separation behavior. In a second piece of work, molecular dynamics was used to calculate diffusion coefficients of small gas molecules in polymer membranes. Other work, performed in collaboration with the Royal Institution, London, involved *ab initio* cluster calculations to develop improved force fields (Bell et al. 1999).

A separate line of work has been on ionic transport in solid oxide fuel cell electrolytes (Khan et al. 1998). Atomistic simulations were used to investigate solubilities and defect structures of dopants in zirconia. Also, molecular dynamics was used to calculate oxygen-ion diffusion coefficients. That work was carried out in collaboration with the University of Surrey.

Despite the small scale of activity, successes have been achieved. At the same time, the methods can sometimes be too easily oversold. Part of the careful development of credibility has been to emphasize experimental verification and to be very clear about the capabilities and limitations of molecular modeling. For example, the speed of MD is a key limitation for many of the studies to date.

Faster, longer calculations are certainly desirable, but if the degree of approximation is acknowledged and taken into account, valuable results can still be achieved.

REFERENCES

Bates, D.R. 1995. *mc2 Magazine* 2(3), Molecular Simulations Inc.; http://www.msi.com/materials/cases/QPA.html.

Bell, R.G., D.R. Bates, and C.R.A. Catlow. 1999. Computer simulation of gas sorption in cation-containing zeolites. Presentation at the Gordon Conference, New Hampshire.

Khan, M.S., M.S. Islam, and D.R. Bates. 1998. Cation doping and oxygen diffusion in zirconia: A combined atomistic simulation and molecular dynamics study. *J. Mater. Chem.* 8, 2299-2307.

O'Connor, D., P. Barnes, D.R. Bates, and D.F. Lander. 1998. A hydration-controlled nano-valve in a zeolite? *Chem. Commun.* 2527-2528. pp. 724-726.

Site: **Center for Atomic-Scale Materials Physics (CAMP Denmark)**
http://www.fysik.dtu.dk/

CAMP-DTU
Technical University of Denmark, Physics Department
Building 307, DK-2800 Lyngby, Denmark
Tel: (45) 45 25 31 75; (45) 45 93 23 99

CAMP-Aarhus
University of Aarhus, Institute of Physics and Astronomy
Nordre Ringgade 1, DK-8000 Aarhus C, Denmark

Date: July 1999

Report author: M. Neurock

Host: Professor Jens K. Nørskov, norskov@fysik.dtu.dk

BACKGROUND

The Center for Atomic-Scale Materials Physics (CAMP) was established in 1993 through the cooperation of the Danish National Research Foundation, The Technical University of Denmark, and The University of Aarhus in order to understand and impact the atomic features that govern material properties. The Aarhus group is internationally recognized for their efforts at following atomic-scale surface processes and their associated kinetics using scanning tunneling microscopy. The group at the Technical University of Denmark in Lyngby is internationally recognized for theoretical, experimental and modeling of surface processes and kinetics for catalysis and material synthesis.

The CAMP team is composed of ten senior scientific staff including Professors Jens K. Nørskov (Director), Flemming Besenbacher (Assistant Director), Karsten Jacobsen (Assistant Director), Hans Skriver, Ib Chorkendorff, Ivan Stensgaard, and Erik Laegsgaard. In addition to the technical staff, there are 3 computing staff members, 2 associated staff, 12 postdoctoral students and 20 PhD students. There have been numerous long-term visitors who have also been affiliated with CAMP over its existence.

The center integrates world-class theoretical and experimental researchers in an effort to tackle important problems in materials development, catalysis and surface science. The center has three main areas of interest: (1) surface dynamics, (2) surface reactions and catalysis, and (3) metallic nano-structures. A general theme of CAMP is the integration of experiment and theory directed toward understanding metallic nano-structures and their properties. They have used periodic density functional theory, LMTO approaches, as well as MD simulation and kinetic Monte Carlo in order examine the physics and chemistry at different length scales. Experimental efforts have included a wide variety of ultra-high vacuum model system studies, as well as microreactor studies.

CAMP (DTU), which is within walking distance to Haldor Topsøe, is very closely connected with the efforts at Topsøe. The collaboration between the groups at DTU, Aarhus and Haldor Topsøe provides a cohesive team that has been able to interrogate heterogeneous catalytic systems across a wide range of length and time scales. The CAMP group has published over 214 papers as well as numerous book chapters since it was formed in 1993. The group has had three papers published in *Nature* and two papers in *Science*. A full list of their papers is given at http://www.fysik.dtu.dk/papers/publications/pub.html.

R&D ACTIVITIES

Surface Dynamics

The Aarhus group has been using STM to examine the dynamic diffusion process of adsorbates on surfaces. They have been one of the first groups to monitor actual atomic diffusion process with the ability to extract elementary activation barriers for diffusion of adatoms and adsorbates on surfaces. DFT calculations from the group at DTU confirmed the experimental findings for Pt diffusion down on Pt(110) 1x2 surface and hydrogen assisted diffusion of platinum. In addition to following surface kinetics, they are also able to closely monitor the evolution of surface structure resulting from sintering, adsorbate ordering and chemical reaction. Some of their current research themes include atom-resolved STM of diffusing adatoms, adsorbate-induced changes in adatom diffusion, adsorbate structure formation, adatom and vacancy cluster dynamics, and steps.

Surface Reactivity and Catalysis

The Aarhus and DTU groups have worked closely together in an effort to understand the important structural and electronic features of metal surfaces that control their reactivity. They have examined the chemisorption and reactivity of various different small molecules important in the chemical process industry with DFT and LMTO ("linear muffin tin orbital") computational methods, and with STM and molecular beam studies. In an effort to understand how to change the surface properties they have actively examined how surface alloys, step edges, promoters and poisons impact surface reactivity. They have been quite successful in building different structure-property relationships that describe chemisorption and reactivity on model surfaces. For example, they have shown that the d-band center projected onto the surface layer of the metal offers a quantitative probe for how small changes in the electronic surface structure due to the compositional and/or geometric structural changes affect surface reactivity.

Their results for alloying nickel with gold demonstrated that unique surface alloys could be formed that could increase the activation barriers for carbon deposition during steam reforming. The results from the modeling and STM studies prompted Haldor Topsøe to attempt to synthesize these alloyed surfaces. A new bimetallic catalyst was found to be considerably better than the currently used Ni catalyst. Patents were being filed on this new surface alloy and its ability to limit graphite formation. The results have been published (Besenbacher et al., 1998).

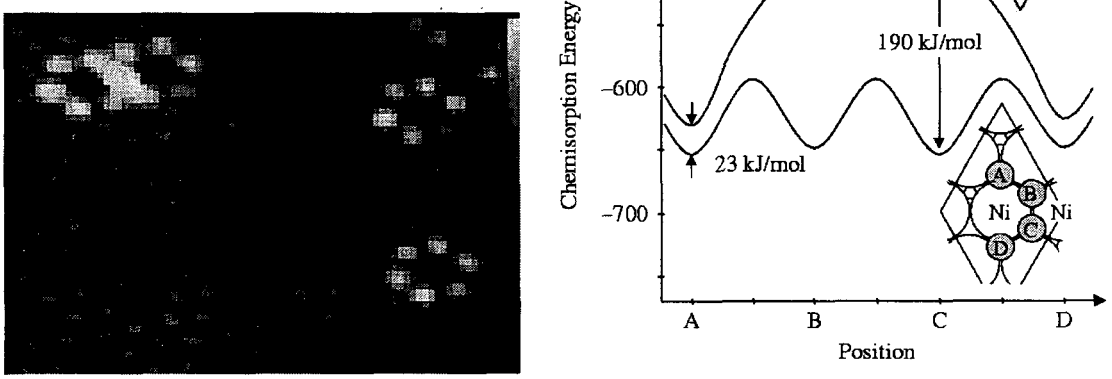


Fig. B1.1. Scanning tunneling microscopy of the Ni/Au alloy (left) and DFT calculations that show the effect of the carbon binding energy upon chemisorption (right).

In addition to the group's efforts on metals, there are smaller but growing efforts to examine the metal sulfide catalysts as well as N_2 activation in FeS enzymes from first-principles efforts.

Other current projects include the following:

- Nanoclusters in HDS catalysis
- Steps dominating N_2 activation on Ru(0001)
- Palladium nanocrystals on Al_2O_3
- High-coverage structures of oxygen on platinum
- Enhanced surface self-diffusion
- Dynamics on metal surfaces
- Point contacts and nanowires
- Reactivity of strained overlayers
- Mechanical properties of nanocrystalline metals
- Solution energies of the 4d metals

Metallic Nanoclusters

CAMP has been actively pursuing how the structure and atomic properties of nano-structures control overall material properties. They have been modeling STM and the conductance in nanowires. In particular they have focused on atomic scale contacts, nanoscale dynamics, nanophase materials, dislocation dynamics and metal-insulator interfaces.

Software Development

In addition to applications of molecular modeling, the group at DTU has been actively developing the CAMP Open Software project (http://www.fysik.dtu.dk/CAMP/CAMPOS_welcome.html). A key part is a planewave pseudopotential code called DACAPO. CAMP currently supports a full-time research staff individual (Lars Hansen) to carry out code development. CAMP recently released a new version of DACAPO that contains a fair number of ultrasoft pseudopotentials for transition metals of interest to heterogeneous catalysis.

In addition to planewave DFT code, the DTU group has developed a "linear muffin tin orbital" method, which has been used to understand the nature of alloying in bimetallic systems. In the early 1980s, Professor Nørskov developed an approximate total energy method that he called the "effective medium theory."

Dr. Ole Nielsen is manager of the computer services group at the Department of Physics and is responsible for maintaining the hardware at CAMP. CAMP maintained a 24-node SP2 machine for five years previous to the WTEC visit. More recently they received a grant from Villum Kann Rasmussen Foundation for approximately $400,000 to configure and install a new Beowulf cluster of Compaq/DEC Alpha workstations (140 nodes as of January 2001).

REFERENCE

Besenbacher, F., I. Chorkendorff, B.S. Clausen, B. Hammer, A. Molenbroek, J.K. Nørskov and I. Steensgaard. 1998. Design of a surface alloy catalyst for steam reforming. *Science* 279, 1913.

Site:	**Centre Européen de Calcul Atomique et Moléculaire (CECAM)** **Ecole Normale Supérieure** **46, Allée d'Italie; Lyon cedex 07** **France** **Tel: +33-4-72728637; Fax: +33-4-72728636;** cecam@cecam.fr **http://www.cecam.fr/**
Date Visited:	16 September 1999
WTEC Attendees:	P.T. Cummings (report author), P. Vashishta
Host:	Professor Michel Mareschal, Director; mmaresch@cecam.fr

BACKGROUND

CECAM (Centre Européen de Calcul Atomique et Moléculaire, http://www.cecam.fr/) is a research center for the advanced study and application of computational science in atomic, molecular, and condensed-matter physics and chemistry. It was founded by Carl Moser in 1969 (thus celebrating its 30th anniversary in 1999) and has played a significant role as a promoter of computational science in the European scientific community. CECAM is based on a collaboration between national research institutions of various European countries, and it is at present hosted by the Ecole Normale Supérieure de Lyon (France).

Funding for CECAM is provided by the following organizations:

- Centre National de la Recherche Scientifique (CNRS, France)
- Commissariat à l'Energie Atomique (CEA, France)
- Consiglio Nazionale delle Ricerche (CNR, Italy)
- Ecole Normale Supérieure de Lyon (ENSL, France)
- Fonds National de la Recherche Scientifique (FNRS, Belgium)
- Fonds National Suisse (FNS, Switzerland)
- Foundation for Research and Technology Hellas (FORTH, Greece)
- Istituto Nazionale di Fisica della Materia (INFM, Italy)
- Nationaal Fonds voor Wetenschappelijk Onderzoek (NFWO, Belgium)
- Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO, Netherlands)
- Engineering and Physical Science Research Council (EPSRC, United Kingdom)
- Deutsche Forschunggemeinshaft (DFG, Germany)—joined in 2001

The research foundations of Spain and Portugal are expected to join as sponsors in the near future. There is no ongoing direct support of CECAM by U.S. funding agencies. However, there is currently in place a 3-year grant at the U.S. National Science Foundation which provides support for the participation of junior U.S.-based researchers (primarily students and post-doctoral researchers) in CECAM workshops. There is no guarantee that funding will be continued beyond the current grant.

Annual funding for CECAM is 3 million French francs (FF) per year. About one-third of the budget is used to support workshops (see below); the remainder supports the administration of CECAM and some research activities located at CECAM.

MOLECULAR MODELING ACTIVITIES

CECAM promotes international cooperation among scientists and scientific institutions, primarily by hosting workshops. A very open process for conducting a workshop at CECAM has resulted in its being continually

a facilitator of simulation research in condensed-matter physics, materials science, statistical mechanics, quantum chemistry, atomic and molecular physics, and the physics and chemistry of biomolecules. About 25 workshops are proposed each year, each of about 3-4 days duration. An advisory board, composed of two members from each supporting organization, selects about 18 workshops for funding each year. A budget of around FF 50,000-60,000 is provided to each workshop. A standard CECAM policy for the use of funds is FF 450/person/day for living expenses and FF 2,000/person for travel by U.S. participants. Despite the fact that CECAM has no official or ongoing funding from a U.S. funding agency, it supports U.S. participation at CECAM and has held workshops with U.S. chairs and/or co-chairs. The result of these activities is that within the molecular modeling and simulation (MMS) community, both in Europe and in the United States, CECAM is highly regarded as a venue for communicating new MMS methodologies, for information exchange between research groups, and for learning about new methods. About 500 researchers attend CECAM workshops each year.

A new activity for CECAM is facilitating tutorials on specific MMS topics. A tutorial consists of a small number of experts in a particular MMS sub-field providing lectures to students (who might be students, post-doctoral researchers, or established researchers wishing to learn) on methods and hands-on experience with their application. This differs from a workshop in that the information exchange is largely uni-directional (from instructor to student) in tutorials, while the workshop format is predominantly for the exchange of information among established researchers.

CECAM also maintains its own research program, the extent of which is largely determined by the effort of the director, who is appointed for a 4-year full-time term. The current director, Michel Mareschal, is on leave from the Université Libre de Bruxelles, and he is currently building up the research activities at CECAM. The only permanent staff member at CECAM is the secretary to the director; other staff members are supported by research grants. Professor Mareschal is also head of the SIMU project (http://simu.ulb.ac.be/home.html). SIMU is the acronym of a research program entitled "Challenges in Molecular Simulations: Bridging the time-scale and length-scale gap" and funded by the European Science Foundation (http://www.esf.org/). The major focus of SIMU is information exchange among MMS research groups in Europe, and it also funds workshops located at CECAM.

CECAM has had a dramatic effect on the propagation of MMS throughout Europe. The first workshop at CECAM brought the developer of molecular dynamics, Anees Rahman, from the United States to enable European researchers to learn the details of how to perform such simulations. Several well-known MMS methods were outgrowths of research performed at CECAM by CECAM post-doctoral researchers or were initiated by CECAM workshops. These include the continuum configurational-bias Monte Carlo method, attributed to Daan Frankel during his CECAM post-doctoral fellowship, and constraint dynamics and the Evans/Gillam method for simulating thermal diffusion in mixtures. As an infrastructural support agency for MMS activities, CECAM has no equal. There is no comparable U.S. activity.

CONCLUSION

CECAM is an extremely valuable resource supporting European MMS activities. It represents a remarkable collaboration by national research organizations to support an activity seen as critical to European science.

Site:	**Centre National de la Recherche Scientifique (CNRS)** **UMR 9986 CNRS-CPE Lyon** **Bât. F308, 43 Bd du 11 Novembre 1918; F-69616** **Villerbanne Cedex** **France**
Date Visited:	16 September 1999
WTEC Attendees:	K.K. Irikura (report author), A.M. Chaka
Hosts:	Prof. Jean Marie Basset, Director, Laboratory of Surface Organometallic Chemistry, CPE (Ecole Supérieure de Chimie Physique Electronique) Lyon Tel.: +33-4-7243-1794; Fax +33-4-7243-1793; basset@coms1.cpe.fr Prof. Philippe Sautet, Laboratory of Theoretical Chemistry, Ecole Normale Supérieure de Lyon

BACKGROUND

CNRS (http://www.cnrs.org/) was created in 1939 by merging all the public institutions for basic and applied scientific research. After World War II, other organizations were created specifically for applied work, and CNRS focused on basic research. In 1966, the organization began providing university research laboratories with both funding and research personnel. CNRS is found throughout French science, much as the Howard Hughes Medical Institute pervades the life sciences in the United States. In the 1970s CNRS embraced applied research once again. In the 1980s a new strategy was adopted to focus upon interdisciplinary programs that address specific areas of societal need. Collaborations involving private industry and other government organizations are also pursued vigorously. At present, CNRS has 25,400 employees, 1,300 service and research units, 3,000 industrial contracts, and a budget of $2.5 billion. Excluding medical research, CNRS laboratories contribute to about two-thirds of all scientific publications in France.

CPE Lyon (Ecole Supérieure de Chimie Physique Electronique de Lyon, http://www.cpe.fr/) was created in 1955 through the merger of the Advanced School of Industrial Chemistry of Lyon (ESCIL) and the Institute of Industrial Chemistry and Physics (ICPI). Its 400 researchers (including 130 PhD students) and 720 students focus on analytical sciences, organic chemistry, process engineering, and information technology. CPE has strong associations with the Claude Bernard University of Lyon and with CNRS and maintains about 100 industrial research contracts. Students are encouraged to spend time studying abroad, in particular doing industrial research.

At least two foreign languages, including English, are required of all students. The Laboratory of Surface Organometallic Chemistry, founded a few years ago by Prof. Basset, is part of the process-engineering department. Its goal is to apply the principles of organometallic chemistry to heterogeneous catalysis.

ENS Lyon (Ecole Normale Supérieure de Lyon, http://www.ens-lyon.fr/) was established in 1987 by merging the science departments of the (women's) ENS de Fontenay and the (men's) ENS de Saint-Cloud. The other departments will also be moved to Lyon in 2000. At ENS Lyon, about 60% of the budget goes to research. The total research budget is more than $7 million, of which about one-third comes from governmental and industrial contracts. There are 460 research staff (including 140 PhD students). The Laboratory of Theoretical Chemistry is directed by Bernard Bigot, who is also ENS's Director for Research and one of the three founders of ENS. The laboratory includes about 25 staff (including several PhD students). It is part of CNRS's Institute of Catalysis Research (IRC), constituting its "Theory and Modeling" theme, of which Prof. Sautet is the director. This research effort is directed at heterogeneous catalysis and includes close collaborations with experimentalists.

R&D ACTIVITIES

Jean-Marie Basset

Professor Jean-Marie Basset has been doing experimental work in catalysis for about 30 years, and was previously part of IRC. There has been little conceptual progress in heterogeneous catalysis since Sabatier in relating structure and reactivity. Most research is merely screening or testing, which is very inefficient. Scientific studies of commercial catalysts are difficult and require very large, expensive instrumentation. Furthermore, the conclusions are often incorrect because the active sites compose such a small fraction of the material.

In contrast, the science of homogeneous catalysis has been spectacularly successful because the full force of modern organometallic chemistry can be applied. The reason for starting this laboratory was to try a new approach: prepare well-characterized heterogeneous catalysts with 100% active sites. This allows the concepts from organometallic chemistry to be used to design catalysts rationally. The materials are prepared by functionalizing oxides, metals, or zeolites using organo-transition metal complexes. The laboratory has 12 major industrial collaborators, including Goodyear, Michelin, Rhône-Poulenc, IFP, BP-Amoco, Hoechst, BASF, Shell, and Elf Atochem, who believe in this strategy. So far, the approach has yielded two major discoveries: catalytic depolymerization of polyolefins and catalytic alkane metathesis. Very recently, they have also discovered a related route for the asymmetric oxidation of olefins. There is some staff in the laboratory doing molecular modeling, but no theoreticians.

Philippe Sautet

For the last ten years or so, Professor Sautet's research has addressed the detailed chemistry of heterogeneous catalysts, chemi- and physisorption, and also the interpretation of surface analytical measurements, especially scanning tunneling microscopy (STM). The theory and modeling group uses many types of theory, ranging from quantum chemistry to statistical physics. This is necessary because they must address not just molecular chemistry, but also porous diffusion, sintering, etc. They have their own Monte Carlo codes, but in general restrict programming to specific applications, such as STM. Most of their quantum calculations are for periodic systems, not isolated molecules. Theirs is the first laboratory in the country that was funded specifically for applications of theory, as opposed to developing theoretical methods. Now there are also such groups in Poitiers and Lille.

The group always works with experimentalists and has acquired a large number of such collaborators. They are also involved with CECAM, are in the GdR, have worked with IFP on hydrodesulfurization, and have collaborated with Rhône-Poulenc on nickel-catalyzed hydrogenation. Nonetheless, they generally do not have strong contacts with industry; such connections are mostly indirect, through their experimental collaborators.

DISCUSSION

The hosts are at the forefront of experimental and theoretical approaches to developing and understanding heterogeneous catalysts. The expressed a number of views during this WTEC visit, as summarized below:

Catalysis is fundamental to the chemical industry and to the broader economy. For example, about 30% of the GDP of the United States is made possible by chemical catalysts.

Modeling is very important but is not used enough. A dramatic, but simple, example is the class of catalysts for converting methane to ethylene at 700°C. Yields were only 20% to 25%, and there were many intensive efforts, totaling billions of dollars, to improve the yield through Edisonian testing. Finally someone examined the supposed mechanism of the reaction and assembled a straightforward kinetic model, using rate constants obtained from combustion research. The modeling showed that the maximum theoretical yield was 25%, and that the massive testing was futile. A tremendous amount of money and time were wasted because the modeling was not done at the outset.

Molecular-level modeling has as much to contribute to catalysis as it has contributed in other areas of chemistry. Theory does not compete with creative chemical intuition, but adds to it by providing an additional perspective. Such perspective may come from sophisticated theory, but more often comes from the simplest models. For example, a picture of a complex molecular structure, or of the physical location of molecular orbitals, will often clarify experimental observations.

Interactions between experimental and theoretical research groups are necessary. Establishing such interactions can be difficult, at least in the area of heterogeneous catalysis. This is largely because the questions are often ill-posed (e.g., catalyst structure not known in atomic detail), extremely difficult (e.g., requiring a PhD thesis in answer), or both. The jargon in the two communities is different and may be an additional barrier. With such problems, even if the theoretician is willing to attack the problem, it will likely take so long that the experimentalist becomes frustrated. It would be helpful to have a database of surfaces to assist with the structural issues. Because strong collaborations are hard to establish, all experimental research laboratories should have at least one theoretician. However, that theoretician risks becoming isolated from technical developments. Sharing a joint PhD student is a good way to collaborate while avoiding such isolation.

Simplified theory should be incorporated into all chemistry courses, perhaps even at the high-school level. On the other hand, in some universities there are faculty who believe that there is too much theory in the courses. All graduate students should be encouraged to use molecular modeling software, so that they become comfortable using modeling to help solve problems. Some years ago there was a wave of new molecular modeling programs in industry, but they were largely discontinued in a subsequent trough of disappointment. Industry is impatient. Now, there seems to be another wave building, with more modeling people hired by industry. One must usually strike a compromise between the chemical complexity of a system and the accuracy obtained from modeling. Fortunately, the point of compromise is moving steadily in the favorable direction, toward greater accuracy for more complex systems. Hopefully this trend will make the second industrial wave of modeling more successful than the first.

CONCLUSIONS

This site shows how experimental and modeling research can complement each other, bringing together a creative tension between the capabilities and limitations of each aspect. Professor Basset's extensive collaborations with industry provide another way that successful experimental/modeling efforts can be communicated.

REFERENCES

Dufaud, V., and J.M. Basset. 1998. *Angew. Chem. Int. Ed. Engl.* 37, 806-810.

Labinger, J.A., and K.C. Ott. 1987. *J. Phys. Chem.* 91, 2682-2684.

Sautet, P. 1997. *Chem. Rev.* 97, 1097-1116.

Vidal, V., A. Théolier, J. Thi volle-Cazat, and J.M. Basset. 1997. *Science* 276, 99-102.

Site: **Daresbury Laboratory**
Computational Science and Engineering Department
Daresbury, Warrington WA4 4AD, UK
http://www.cse.dl.ac.uk/

Date Visited: 14 September 1999

WTEC Attendees: P. Vashishta (report author), A.M. Chaka, P.T. Cummings, K.K. Irikura, P.R. Westmoreland

Host: Dr. Paul Durham, p.durham@dl.ac.uk
Prof. Nicholas M. Harrison (Coordinating host), n.harrison@dl.ac.uk
Dr. William Smith, w.smith@dl.ac.uk
Dr. Paul Sherwood, p.sherwood@dl.ac.uk
Dr. Martyn Guest, m.guest@dl.ac.uk
Dr. Richard Blake, r.j.blake@dl.ac.uk

BACKGROUND

Dr. Paul Durham, Head of the Computational Science and Engineering Department, welcomed the WTEC group. Dr. Durham conducted the meeting along with Prof. Nicholas Harrison, who was the contact person for the visit. Dr. Durham described the background of Daresbury Laboratory and the organization of research areas. Drs. Durham, Smith, Sherwood, Harrison, Guest, and Blake, who all belong to the Computational Science and Engineering Department, made scientific presentations.

COMPUTATIONAL SCIENCE AND ENGINEERING AT DARESBURY

At Daresbury, there are 50 computational scientists. Research councils provide about 80% of the funding and the remaining 20% comes from industry and European Union sources. Daresbury scientists wrote many of the basic codes used in UK research, and the CSE group supports a large fraction of application codes used by the UK community. The users of CSE-developed codes and services can be divided into four categories as shown in Table B1.2.

Table B1.2
CSE-developed Codes and Services

Data services (CDS)	>1900 users
Cooperative Computational Projects (CCPs)	~200 groups
High Performance Computing Initiative (HPCI)	~150 groups
	~1000 users
European Union (EU) Networks	~200 users (50 groups in EU)

The Council for Central Laboratory of the Research Councils (CCLRC) plays an important role in computational science research and infrastructure development. There are four aspects of CCLRC involvement in computational sciences:

- Academic research—theory and computation (in conjunction with experiment) and grand-challenge simulations
- Industrial research—in collaboration with industrial research laboratories and contract research
- International collaborations
- Interaction and collaboration with hardware and software suppliers

There are a number of diverse mechanisms employed to achieve these goals: Collaborative Computational Projects (CCPs); HPCI center; community clubs; industrial collaborations; development facilities; and computer codes.

Collaborative Computational Projects have played a very important role in a broad range of topics in science and engineering. Each project has a chair and a working group, which sets the scientific agenda and decides the work program and monitors progress. CCPs bring together the major groups to tackle large software development, maintenance and distribution projects. Projects and leaders are shown in Table B1.3.

Table B1.3
Collaborative Computational Projects

CCP1	Quantum Chemistry: The Electronic Structure of Molecules (http://www.cse.clrc.ac.uk/Activity/CCP1)	Professor M. Robb
CCP2	Atomic & Molecular Physics: Continuum States of Atoms and Molecules (http://www.dl.ac.uk/CCP/CCP2/main.html)	Professor K. Taylor
CCP3	Surface Science: Computer Simulation of Surface Structure and Properties (http://www.dl.ac.uk/CCP/CCP3/main.html)	Dr. A. Fisher
CCP4	Protein Crystallography (http://www.dl.ac.uk/CCP/CCP4/main.html)	Professor N. Isaacs
CCP5	Molecular Simulations: Computer Simulation by Molecular Dynamics, Monte Carlo and Molecular Mechanics of Liquid and Solid Phases (http://www.dl.ac.uk/CCP/CCP5/main.html)	Prof. D. Hayes
CCP6	Heavy Particle Dynamics	Professor J. Hutson
CCP7	Astronomical Spectra	Professor D. Flower
CCP9	Electronic Structure of Solids	Dr. J. Annett
CCP11	Biosequences and Function	Professor D. Moss
CCP12	Computational Engineering	Dr. S. Cant
CCP13	Fiber Diffraction	Dr. T. Forsythe
CCP14	Powder and Single Crystal Diffraction	Dr. J. Cockcroft

Typical activities include the following:

- Implement "flagship" code development projects
- Maintain and distribute code libraries
- Organize training in the use of codes
- Hold meetings and workshops
- Invite overseas researchers for lecture tours and collaborative visits
- Issue regular newsletters

Current CCPs involve about 240 UK academic groups in most UK universities, plus several research institutes and industrial laboratories. The CCPs also have about 150 collaborating groups in Europe, the United States, other countries, and in EU Networks. The CCPs are funded by EPSRC, BBSRC, and PPARC.

The CCPs provide a framework within which all the active groups can work together. Their achievements include the following:

- Very large number (>300) of codes supported on many hardware platforms
- High-quality publication record (the projects produce an average of about 64 publications/year each, of which about 11% are in the *Physical Review Letters* and *Nature* "hot news" category)
- Strong presence in "Grand Challenge" consortia within the HPCL
- Extensive participation in European programs (HCM, TMR, ESF, Networks)
- Strong interest from industrial research groups
- Longevity—funding maintained within rigorous peer-review process

One of the important functions of the Computational Science and Engineering Department is to provide computational development and application support. Application support encompasses the following:

- Development of codes with new functionality
- Harnessing relevant developments within the computer science and numerical analysis arena
- Performance evaluation and optimization
- Software maintenance and distribution
- Training and information dissemination

It is hoped that application support will foster new, especially interdisciplinary, computational areas. To achieve this goal, a broad range of scientific expertise is needed in one place. A key point is that when collaborative development works well, then computational developments make a rapid impact in real applications.

International Collaborations

Daresbury scientists have broad-based international collaborations with European research groups and national laboratories in the United States and Japan.

- Europe: Several EU and ESF-funded networks in solid state physics and chemistry, atomic physics, and materials science; ESPRIT projects in computational fluid dynamics, molecular simulation and quantum chemistry; member of ERCIM and ERCOFTAC.
- U.S.: Close collaboration with the Department of Energy laboratories and National Institutes of Health. Among the DOE laboratories there is close collaboration with Pacific Northwest and Oak Ridge and Lawrence Livermore National Laboratories. The interaction with Pacific Northwest is especially strong because Dr. Guest was on the research staff at the laboratory for several years. Other Daresbury scientists have also visited and collaborated with researchers at DOE laboratories.
- Japan: Daresbury scientists collaborate with RIKEN (a national laboratory in Japan; see separate site report) in the areas of chemistry and materials science. There is also collaboration with CSSC on software development for massively parallel processor systems.

Industrial Collaboration and High-Performance Computing

Daresbury scientists are systematically and deliberately increasing their collaboration with a number of industries in the UK. The computational science and engineering department has a substantial program of industrial collaboration with ICI, Zeneca, Unilever, Shell, DRA (Defense Research Establishment), and BAE (British Aerospace Engineering). There is also close collaboration with computer companies (Cray/SGI, IBM, Compaq, etc.) and software houses such as MSI. The laboratory has broad-based strength in the area of high-performance computing and its applications to modeling and simulations that make it very attractive for industrial research laboratories to interact and collaborate with Daresbury scientists.

Industrial research laboratories are interested in collaborating with Daresbury because of the following strengths in its computational science and engineering department:

- The CSE department understands the growing potential of computational sciences.
- Daresbury scientists have experience in simulation methods and algorithms. They can assist simulation and modeling scientists in industry by quantifying their computational requirements and help them make the transition from workstation to high-performance computing capability.
- Daresbury scientists can collaborate with industrial researchers to solve hitherto intractable problems.
- Industrial researchers can gain cost-effective and low-risk access to the high ground of computational science and engineering.

There are further plans for two kinds of industrial collaborations:

- Daresbury scientists are planning to build a consortium of researchers from computational science and engineering departments, academic institutions, industrial researchers, and hardware/software suppliers.
- Daresbury scientists are also targeting industry for funding to support high-performance research and development activities. This kind of funding will benefit all researchers—academic as well as industrial—who use high-performance computing resources at Daresbury.

There are many barriers to and difficulties in working with industry. Some of these difficulties are inherent in industrial research whereas other problems can be overcome by mutual education and increased interactions between Daresbury scientists and industrial researchers. Examples of difficulties encountered in collaborating with industry by Daresbury scientists and academic researchers include: short time scales of industrial research; intellectual property rights and security; industry's lack of awareness of the cost-benefit of computational modeling; and (self-critically) lack of awareness of what industry's real problems are by Daresbury and academic researchers.

There is a strong need for a continuous dialogue and mutual education about each others' concerns for interesting subjects to be tackled by Daresbury-industry collaboration.

Future Plans

Four major areas were outlined for future advances:

- HPC200x—Assist EPSRC with the scientific case for computational science, modeling and simulation; novel hardware approaches (Beowulf clusters, etc.)
- Using the Web for novel delivery of integrated services
- Integrated modeling on many length scales: microfluidics, environmental modeling
- Computational biology: Modeling of organelles, cells and their interactions

MODELING R&D ACTIVITIES

After a presentation of the organizational overview and objectives by Dr. Paul Durham, five research presentations (Drs. Bill Smith, Paul Sherwood, Nic Harrison, Martyn Guest, and Richard Blake) were made to the panel members. These scientists come from diverse backgrounds in solid-state physics and quantum chemistry. Common features of nearly of all the modeling and simulation research discussed were that (1) the projects involved both simulation and some code development and (2) there was some contact with relevant industrial scientists and academic researchers.

Dr. William Smith described research on molecular simulation of materials using DL_POLY, the molecular dynamics program developed at Daresbury during last several years. Under the DL_POLY project, a general-purpose parallel MD code was developed in 1992 with the support of CCP5/EPSRC projects. The MD code is quite versatile and has the following features: Ewald sums for long range interactions; shell model and distributed multipoles; NVE, NVT and NPT ensembles and their variations; flexible and rigid bond constraints implements using SHAKE algorithm; 2-, 3-, 4- and N-body non-bonded forces; and parallel

and serial operation. The parallel implementation until 1999 used replicated data on each node. This was a severe limitation for large simulations. Typically up to 30,000 atoms on 100 processors could be done using replicated data methods. Domain decomposition was incorporated in DL_POLY in 1999. This allows for million atom simulations using link cells and permits o(N) atom scaling and o(P) processor scaling. The electrostatics is handled using SPME (smooth particle mesh Ewald) method. A variety of applications have been implemented using the DL_POLY code. These include ionic systems, solutions, biological systems, metals, and polymers. Simulation results for the structure and alkali diffusion for $M_2Si_2O_5$ silicate system and K^+ complexation in valinomycin using the AMBER force field were presented.

Dr. Paul Sherwood discussed quantum chemistry research. At Daresbury, quantum chemistry is one of the strongest areas. This included high performance computing software development and applications of the quantum chemistry tools to QUASI collaboration.

The HPC software development includes the following:

- Quantum Chemistry (GAMESS-UK and NWChem)
- Integrated Modeling Schemes (ChemShell and GAMESS-UK/CHARMM)

GAMESS-UK (http://www.dl.ac.uk/TCSC/QuantumChem/Codes/GAMESS-UK/) is a general-purpose *ab initio* molecular electronic structure program for performing SCF-, MCSCF- and DFT-gradient calculations; it also implements a variety of post-Hartree-Fock techniques. The program is derived from the original GAMESS code, obtained from Michel Dupuis in 1981 (then at the National Resource for Computation in Chemistry, NRCC, at Lawrence Berkeley Laboratory), and it has been extensively modified and enhanced over past decade. The development of GAMESS-UK has included contributions from numerous authors (M.F. Guest, J.H. van Lenthe, J. Kendrick, K. Schoffel and P. Sherwood, with contributions also from R.D. Amos, R.J. Buenker, H.H. van Dam, M. Dupuis, N.C. Handy, I.H. Hillier, P.J. Knowles, V. Bonacic-Koutecky, W. von Niessen, R.J. Harrison, A.P. Rendell, V.R. Saunders, and A.J. Stone). The work has been conducted largely at CCLRC Daresbury, under the auspices of Collaborative Computational Project number 1 (CCP1). Other major supporters that have assisted in the ongoing development and support of the program include various academic funding agencies in the Netherlands. Industrial support has come from ICI plc. GAMESS-UK is marketed by the company Computing for Science Ltd. (http://www.dl.ac.uk/CFS), which was founded in 1992 by a consortium led by Martyn Guest (Daresbury), J.H. van Lenthe (Utrecht), J. Kendrick (ICI), and K. Schoffel (Norsk-Hydro).

Development of NWChem software (http://www.emsl.pnl.gov:2080/docs/nwchem/) is being led by the Environmental Molecular Science Laboratory (EMSL)—a U.S. Department of Energy laboratory located in Richland, Washington. The EMSL was funded by DOE to be an integral component in solving grand-challenge environmental restoration problems. NWChem was designed and developed to be a highly efficient and portable MPP computational chemistry package. The software provides computational chemistry solutions, which are scalable with respect to chemical system size as well as MPP hardware size.

Integrated approaches combining quantum mechanics (first-principles energy expression based on electronic and nuclear coordinates) and molecular mechanics (cost-effective parametric energy expression based on nuclear positions) are also being developed. These QM/MM integrated approaches allow QM computations in the active region of the system and MM for the region surrounding the active region.

Quantum Simulation in Industry (QUASI), a multi-institution partnership, was described. Members include the following organizations:

- CLRC Daresbury Laboratory (Drs. P. Sherwood, M.F. Guest, and A.H. de Vries)
- Royal Institution of Great Britain (Prof. C.R.A. Catlow and Dr. Alexei Sokol)
- University of Zurich (Switzerland)/MPI Mulheim (Germany) (Prof. Walter Thiel, Drs. Salomon Billeter, Frank Terstengen and Alex Turner)
- ICI Wilton (UK) (Drs. John Kendrick (CAPS group) and John Casci (Catalco))

- Norsk Hydro (Porsgrunn, Norway) (Drs. Klaus Schoeffel and Ole Swang, SINTEF)
- BASF (Ludwigshafen, Germany) (Dr. Ansgar Schaefer)

A number of applications have been implemented in this collaborative project. Examples are enzyme catalysis applications with BASF as the lead partner; surface catalysis applications with ICI as the lead partner; and zeolite catalysis applications with Norsk-Hydro as the lead partner.

Prof. Nicholas Harrison described research in computational materials science. This is a very high-quality program involving first-principles simulation (energy, free energy, dynamics, electronic structure, and magnetism), analytic instruments (diffraction and spectroscopy), and visualization.

This effort is well-anchored in four directions: (1) with CCP3 (surface science development) on the computational front; (2) with experimental facilities (SRS, ISIS, and MEIS); (3) with an active partnership involving MSI for code development and high-performance computing software using plane-wave density functional theory; and (4) with collaboration with industry to attack grand challenge applications of mutual interest. A number of computer codes are used in this effort. All these codes have been developed at Daresbury or in collaboration with Daresbury scientists:

- CRYSTAL: Periodic LCAO DFT and Hartree-Fock, 1974 - present, Dovesi, Roetti (Turin) Saunders, Harrison (Daresbury)
- CASTEP: Plane-wave, pseudo-potential DFT, 1988 - present, UK Car-Parrinello consortium / MSI
- DL-LEED, DL-EXCURVE, DL-XANES, PHOTON: Greens function multiple scattering approach, CCP3, SRRTNet
- DLVisualize: Environment and GUI for materials modeling

One of the problems studied in great detail involves TiO_2 surface with water. Structure of the surface without water and with different coverage of water are computed using first-principles codes. The vibrational spectra are also computed and compared to experimental data to resolve a number of outstanding issues related to dissociation of water on the surface. In the next stage of code development, issues related to the excited states, thermodynamic functions at finite temperatures, and overall ease of use of the software tool on MPP platforms will be addressed.

Dr. Martyn Guest described the status of computational chemistry on high-end and commodity-type computers. He is a very experienced computational chemist who was a research scientist at the U.S. DOE Pacific Northwest National Laboratory before moving to Daresbury a few years ago. He has a broad perspective and is aware of strengths and weaknesses of the systems and approaches in the United States and in the UK. He discussed a broad range of topics, as listed below:

- High-performance computing in the UK (opportunities and current investment levels; issues of cost-effectiveness: high-end, departmental and desktop)
- High-performance computing and CLRC (Daresbury's HPCI Center and support for distributed computing; scalable & distributed high-performance computing facility)
- High-end computational chemistry (GAMESS-UK and DL_POLY—CSAR T3E/1200E)
- Computational chemistry on commodity supercomputers
- Future objectives and HPC requirements (performance modeling on GAMESS-UK and DL_POLY)

The opportunities and current investment levels in UK academic high-performance computing are as follows:

- Decisions on top-end machine (576 node T3E/1200E) and its successor
  - Six-year PFI programme valued at £26+M
  - CSC/Manchester-based consortium (SGI/Cray T3E)
  - Research Councils (HSC) currently deciding on start-date for HPC2x

- Mid-range systems dominate funding streams
  - Joint Research Equipment Initiative (JREI)—HEFCE 3-year investment in mid-range HPC systems (8-32 processor systems) totals some £10-15M
  - Joint Infrastructure Fund (JIF), where JIF'99 bids were under consideration for £25M in university-based mid-range systems; applications dominated by SGI hardware
  - EPSRC's Multi-Project Equipment Initiative

Dr. Richard Blake is the coordinator of industrial initiatives. He discussed industrial applications and integrated modeling carried out by Daresbury scientists in collaboration/cooperation with industrial research scientists. He gave a list of industrial partners (Table B1.4) and discussed various modes of industry and Daresbury collaboration.

Table B1.4
Industrial partners of Daresbury's Computational Science and Engineering Department.

Industrial End-Users	Computer Vendors	Software Companies
ICI Katalco	IBM—marketing, cluster computing, HPC 97	MSI
ICI Tioxide	Fujitsu—HPC 97	Oxford Materials Ltd.
Eutech	HP—cluster computing	Rockfield Software
CAPS	Cray—Research Fellowships	FEGS Ltd.
Zeneca	Compaq (& QSW)	NAG Ltd.
British Aerospace	Numerous developers of HPC Beowulf solutions	PALLAS
DERA Farnborough		Abaqus
Rolls Royce		LS-Dyna
British Steel (now Corus)		STAR-CD
GEC Alsthom		Fluent
Farside Technologies		AEA Harwell
Paras		
Bertin		
Labein		
Electricidade de Portugal		
Iberdrola		
Mitsui Babcock Energy Limited		
Philips		
Siemens		
Inco Alloys Ltd.		
Thermotech Ltd.		
Unilever		
Shell		
Courtaulds		
BP		
DuPont		
ABB Ltd.		
Norsk Hydro		
BASF		
Saint Gobain		
Indo		
Pilkington		
Cemef		
CEA		

He noted various ways of working with these partners:

- Consultancy—HPC technology, scientific developments and opportunities, corporate HPC requirements, procurement support, business cases, modeling technology
- Education—HPC hardware and software technology and scientific developments
- Facilities—cost-effective, secure, low-risk access to medium and large-scale HPC facilities, well-supported by a quality user-service running in-house
- Training—in the use of codes and facilities and implementation on advanced technology
- Technology implementation—in customer, academic or commercial codes
- Demonstrations—prototype calculations of technology or science
- Commercial software computing environment—exploitation of well-packaged and -supported, leading-edge academic applications and commercial packages on central HPC facilities
- Industrial grand-challenge projects—full research collaborations aimed at solving key medium-term industrial R&D problems
- Commercial software development—access to large-scale project funding schemes for longer-term applications development/validation/demonstration/commercialisation (LINK, Foresight, CEC)

FUTURE OF COMPUTATIONAL MATERIALS SCIENCE AND MATERIALS SIMULATIONS

Daresbury scientists expect that there will be continuing demand for access to highest-performance systems to extend accuracy and scope of the current academic applications base. These will include evolution of existing applications:

- Molecular processes on surfaces, design of materials with specific properties and functions—meso-scale modeling
- Predictive drug design and delivery mechanisms
- Realistic turbulence models, rheological properties of materials, microfluidics and MEMS technology

Applications that have yet to realize benefits may include the following:

- Computational mechanics, device simulation, hydraulics, bio-engineering, network distribution, marine CFD, geophysics
- Use of simulation in real-time and time-critical applications, non-numeric applications, IT applications

The Daresbury scientists articulate the 2000-2010 target for computational science as "Towards Petaflops Computing"—a decade of theory/algorithm development. Many problems in materials simulations will need this computing power. Some of the problems are enumerated below:

- High-consequence problems—catalysis, combustion, environmental
- Realizing the "vision" of interactive design and hypothesis testing (PSEs)
- Desk-side/department teraflop computing will satisfy many applications
- How big will be the pool of supercomputer users?
- Basic research is necessary
- Theoretical and computational chemistry, computer science, applied math cannot be isolated from each other—lesson from HPCI—need a focus
- Serious shortage of skilled people with multidisciplinary training and lack of collaboration between computational physicists/chemists and computer scientists—timelines and expectations must take this into account
- Performance, functionality, PSEs (data browsers, advisors.), training, etc.
- Barriers for non-experts (long start-up times)

CONCLUSION

Daresbury continues to be at the heart of many significant developments in applied computational science and in molecular and materials modeling. Their achievements have been due both to intellect and also to their commitments to high-quality research and to collaborations—with the industrial community, with academia, and with the international community.

Site: **Degussa-Hüls AG (now Degussa)**
Hüls Infracor GmbH
Paul-Baumann Strasse 1
Marl, 45764
Germany
http://www.degussa-huels.de

Date Visited: 14 September 1999

WTEC Attendees: E. Stechel (report author), P. Kollman, R. Miranda, K. Morokuma

Host: Robert Franke (49-2365-49-9892), robert.franke@infracor.de

BACKGROUND

Degussa-Hüls is an international company that resulted from the February 1999 merger of two established firms. The combined company started with more than DM 20 billion in sales and a workforce of around 46,000. It is one of the world's largest specialty chemicals companies. Degussa-Hüls AG is run as a decentralized group, with strategic business units or independent subsidiaries responsible for operations. The group comprises 16 strategic business units grouped in four reporting segments:

- The Health and Nutrition segment is one of the leading suppliers of feed additives, principally essential amino acids to improve the nutritional value of the protein contained in animal feed.
- The Specialty Products core business includes colorants, coating raw materials, engineering plastics, bleaching chemicals for the pulp and textile industries, detergent ingredients, and basic intermediates for agricultural and pharmaceutical applications.
- The Polymers and Intermediates segment is the world's second largest producer of methylmethacrylates and polymethylmethacrylates. It includes the Plexiglas® product range.
- The Performance Materials segment's products include automotive catalysts, precious metals, advanced fillers, rubber chemicals, ceramic stains, ceramic colors, decorating and glass colors, glazes, and organic decorating additives.

In addition to the decentralized strategic business units, Degussa-Hüls AG also has centralized entities that provide services for the entire group:

- Degussa Bank GmbH provides innovative financial services.
- Infracor GmbH provides technical and analytical services and operates the infrastructure at the Marl and Hanau-Wolfgang sites.
- CREAVIS GmbH acts as a "venture capital company" responsible for developing new businesses.

(Note: Subsequent to the WTEC visit, in 2000 Degussa-Hüls began another merger, this time with the German chemical company SKW. The direct cause was that their parent companies Veba and Viag were merging into E.ON, a power and utilities company. The name for the combined chemical company is Degussa. In April 2001, Degussa refocused on specialty chemicals, selling several its businesses. Its catalyst business will be owned by OM Group of Cleveland, OH (http://www.omgi.com/), and its ceramics, pigments, and electronic-materials business will be owned by Ferro Corporation (http://www.ferro.com/), also of Cleveland.)

MODELING R&D ACTIVITIES

Computational molecular-scale modeling at Degussa-Hüls is done by two people, one each from the two parent companies. Robert Franke, the panel's host, came from Degussa, although he was hired no more than four months before the merger. He is located at the Marl site. Bernd Hannebauer joined Hüls also only a few months prior to the merger. He was previously at MSI and is now located at the Hanau site.

Computational modeling is part of the centralized unit, Infracor, and as consequence, it operates as a service-for-a-fee to all business units, similar to analytical services. Proximity largely determines which customers will use each of these two computational modelers.

Franke has a 5-node Linux cluster and an SGI Octane. The budget for software is about $50 thousand per year. They have Cerius, InsightII, Turbomol, and DMol, and they plan to purchase Gaussian software.

The expectation is that the computational modeling effort will be self-supporting and profitable by year 2001. Franke is realistic and optimistic about the challenges that these expectations present. He will have to market his expertise to the business units, while simultaneously expanding his capability base. The effort will be evaluated based on the number of customers and the price customers are willing to pay. He cannot turn away projects, although he can and must prioritize his activities. His vision is to eventually be well integrated into the research units, with a science-based effort leading to discoveries, and therefore patents. He expects that computational modeling and modelers should be part of project teams and should be generating ideas and testing hypotheses.

Modeling is presently focused on homogeneous catalysis, computation and correlation of thermochemistry and thermophysical properties, and finding and optimizing leads.

In an earlier success, Alexander Bünz and colleagues from Hüls Infracor at Marl had applied QSPR techniques to correlate and predict normal boiling points for chlorosilanes and to correlate critical temperatures, critical pressures and acentric factors of alcohols (Bünz et al. 1998a; Bünz et al. 1998b). Using data from DIPPR, the Dechema Detherm Databank, the Dortmunder databank, and in-house data, they modeled molecules with Cerius2 (MSI) and fitted QSPR relations with CODESSA (Semichem). For the chlorosilane boiling points, the best descriptors were molecular composition and structure. For the alcohol properties, the most effective descriptors included topological features (for T_c, P_c, acentric factor), constitutive features (T_c, P_c), electrostatic features (P_c, acentric factor), and quantum-chemical features (T_c, acentric factor).

At present, there is not the time or need for outside interactions beyond being a member of the MSI polymer consortium. The primary objective is to get the software, not to benefit from the science.

Breakthroughs that could have large impact would include:

- More efficient density functional algorithms, such as linear scaling
- More reliable functionals for density functional calculations
- More reliable density functional approaches to transition metals

CONCLUSIONS

The computational modeling effort at Degussa-Hüls has a number of obstacles to overcome before it can hope to be considered a sustainable effort. The combination of a merger only seven months prior to the panel's visit and an effort that is in an early stage of development makes it difficult to assess the prospects. However, the company has made a strong initial commitment to this kind of work.

REFERENCES

Bünz, A.P., B. Braun, and R. Janowsky. 1998a. Application of quantitative structure-performance relationship and neural network models for the prediction of physical properties from molecular structure. *Ind. Eng. Chem. Res.* 37(8), 3043.

Bünz, A.P., J. Husung, C. Schröer, and R. Janowsky. 1998b. Quantitative structure-property relationship methods for thermophysical properties of industrial relevance. Topical Conference on Applying Molecular Modeling and Computational Chemistry, American Institute of Chemical Engineers, Miami Beach FL, November 15-20.

Site: **DSM Research**
P.O. Box 18, 6160 MD
Geleen
Netherlands

Date Visited: 17 September 1999

WTEC Attendees: P.R. Westmoreland (report author), A.M. Chaka, K.K. Irikura, P. Vashishta

Hosts: Prof. Dr. Ludo A.L. Kleintjens, Academic Contacts & External Research (http://www.dsm.nl/acer) P.O. Box 18, 6160 MD Geleen; The Netherlands Tel: +31 46-4761938; Fax: +31 46-4761200; ludo.kleintjens@dsm-group.com
Prof. Dr. Robert J. Meier, Research Fellow; Dept. of Physical, Analytical, and Computational Chemistry; r.meier@wxs.nl

BACKGROUND

DSM (http://www.dsm.nl/) is a global company, employing 23,000 people worldwide with about 11,000 in the Netherlands. Sales for 1998 were $7.5 billion. Its products include the following:

- Polymers and industrial chemicals—polyethylenes and polypropylenes (each about 12% of the western European market), ethylene and steam-cracking byproducts for captive use and sale, caprolactam (30% of the world merchant market, world's largest share), acrylonitrile (12% of European market), melamine (30% of world market), and ammonia and nitrate-based fertilizers
- Performance materials—synthetic rubber, ultra-strong polyethylene fibers, nylon, various thermoplastics, polymeric coatings and films, and polymer additives
- Life sciences products—chemical intermediates and active ingredients for pharmaceutical and crop protection industries, antibiotics and antimycotics, and food ingredients and additives

Its stated corporate strategy is to reinforce its strong positions in polymers and industrial chemicals while looking to growth in the life science and performance materials areas. To those ends, about 2,200 people and $250 million are involved in DSM Research, about half in Geleen. DSM also has about 200 research projects at universities (roughly equally split among Dutch, other European universities and the rest of the world). That research totals about $10 million, including about $5 million in large consortia.

R&D ACTIVITIES

A wide range of molecular and materials modeling was described at DSM, including: application to homogeneous catalysis; evaluation of thermodynamic properties; small molecule properties; property prediction for polymers and blends; and interpretation of analytical data. Computational quantum chemistry, molecular simulations, and information management approaches are all applied. This sort of computational molecular-based modeling began at DSM in 1986, although classical polymers modeling and estimation of properties by group contributions were active much earlier.

A core group of eight specialists and technicians focuses on molecular and materials modeling in DSM Research's Department of Physical, Analytical, and Computational Chemistry: two doing molecular modeling, two in polymer modeling, two in expert systems including statistics, and two using computational fluid dynamics. Molecular modeling users are also beginning to be distributed within the company as well, mostly bench-scale chemists using molecule construction and visualization. Past applications and projections are shown in the Table B1.5 below.

Career paths of the hosts illustrate the mix of backgrounds and approaches. Dr. Kleintjens is presently director of external relations (Academic Contacts and External Research). He came to his present job from the technological side, moving from experimental polymer characterization to head the thermodynamics

group that hit a peak size of 15 people in 1985. The group did modeling of high-pressure experiments, of polymer blends using Flory-Huggins approaches, and predictions of separation process parameters for engineering. By contrast, Dr. Meier came from academic backgrounds in theoretical chemistry and experimental physics, beginning at DSM in 1985 with work in spectroscopy and introducing theoretical chemistry. He began doing *ab initio* calculations in 1986 and visualization in 1988. In his role as group leader, he and his colleagues began finding more and more productive uses of computational tools. They quickly began to use Car-Parrinello codes for a range of problems, collaborating on a variety of projects. Currently, Dr. Meier holds a research fellow position with special attention for modeling and spectroscopy and its application in a broad sense.

Table B1.5
Technical Summary of Molecular and Materials Modeling at DSM

Past and Present Applications	Important Emerging Applications	Theoretical Approaches and Software	Hardware
Modeling of homogeneous catalysts (structure and reactivity) (Meier et al. 1994; Aagaard and Meier 1998; Meier 1999) Calculation of thermochemical data (Barrett and Meier 1996) Small molecule (chemicals) properties (Coussens et al. 1992; Meier 1993) Modeling of polymers and blends (phase behavior, mechanical strength, and ultimate properties) (Hageman et al. 1997; Wendling et al. 1999; Meier and Struik 1998; Terizis et al. N.D.a; Terizis et al. N.D.b; Jahromi et al. 1998) Modeling of interpretation of spectroscopic analytical data (Meier et al. 1995; Sleigh et al. 1996; Tarazona et al. 1997; Reinhold et al. 1998; Koglin and Meier 1999)	Computational thermochemistry for process design Heterogeneous catalysis Predicting complex morphologies in polymers using first-principles properties in coarse-grain multi-scale models Using predictions of structural spectra in a partnership with synthesis, materials developers, and analytical services	Electronic density-functional theory, wavefunction-based *ab initio*, some semi-empirical QM, atomistic simulation of solids, data mining, classical modeling with group contributions and polymer theory, mesoscale modeling MSI codes, Car-Parrinello code, Gaussian 98, SPARTAN, Turbomole, CAChe, some internally developed codes	Workstations and PCs moving toward parallel computation on clustered PCs

From the hosts' point of view, a key feature of molecular modeling—and research in general at DSM—is the integration of modeling into DSM's project-based approach. Line management is through various departments, but work is normally carried out in project teams. When projects are formed, the leader selects members with the different needed areas of expertise from within the company. Dr. Kleintjens went so far as to say that in DSM's recruiting, the other most important personal attribute along with technical ability was a collegial attitude toward "sociable interaction." New PhDs initially are usually assigned to DSM Research, where they also learn about the technical resources there before moving out to business units. Accordingly, the molecular/materials modeling work presently comes both from team leader requests and from the group's learning of project opportunities. DSM Research's formal organizational structure was to change January 1, 2000. However, the project approach was to be maintained.

DISCUSSION

For application to catalysis, the perception is that reliability and accuracy of the molecular modeling are not yet sufficiently understood. Calculations made with current methods may reach sufficient maturity in the next 5 to 10 years, and new methods should emerge in that timeframe to be evaluated.

Dr. Meier believes that codes will have to be moved quickly to PC-based platforms, both in high-powered workstations and in parallel configurations. Existing computing platforms, including dedicated parallel supercomputers, may be adequate for the smaller, specific systems often studied in academia, but developing greater industrial value requires not only large-system calculations but also many cases and systems. Rapid development of PC platforms augurs well for their being the best, most economic alternative.

Improvements at each scale, as well as multi-scale techniques, will be important. Larger quantum-chemistry calculations are being applied to many more systems of real-life interest, avoiding the simplifications necessary for classical simulations. Multi-scale methods using scale-spanning within the methods will be increasingly important, but so will be scale-linked calculations—combinations of steps involving different methodologies, such as using the results of quantum calculations for coarse-grained Monte Carlo calculations, in turn generating numbers for design by finite-element methods. Along those line, major changes that may come about due to hardware and software advances in computing power include "infiltration" of advances in computational chemistry into finite-element methods and vice versa. Crack propagation and its role in materials properties is one such example.

Meier noted three issues frequently separating industrial and academic viewpoints:

1. Need for multi-disciplinary approach.
2. The problem-solving approach. Scientific/technological problems require a different approach from more purely intellectual problems, not the least of which is the team approach. DSM management puts great stock in drawing together necessary expertise and operating with people in a team of expertise to bring about faster, more certain solutions to problems that are often ill-defined and open-ended. This approach is quite different from the individual-based experience of graduate or post-doctoral research.
3. Academic focus on developing new methods at the expense of testing and validation. By itself, developing new methods is a crucially valuable role. Unfortunately, many methods are subjected to testing only on a few small problems. Reliability is consequently a routine problem. Fundamentally, running one application after another is not as interesting for academics or as highly regarded.

As an aside, Meier noted "commercial software companies often do far too little on validation;" he fears that the emphasis on developing new customers has often been at the price of de-emphasizing the longer-term development and validation that keeps customers. Jointly funded industry-government validation seems sensible. Again, though, sticking to extremely well characterized species, as in the G1/G2/G3 data sets, neglects the larger organics and inorganics that are frequently of interest in the chemical process industries.

Another role for government would be to finance exchanges of industrial and academic scientists, at faculty, post-doctoral and PhD levels. There are good examples of this approach, but relatively few. A danger is that over-emphasis on technological application by industry and government could kill fundamental science. Company sponsorship of academic research is another mutually valuable approach, which DSM uses heavily.

DMS researchers' biggest desires are the ability to handle bigger problems, to model transition metals more accurately, and to obtain force fields that are more accurate and whose range of accuracy is better characterized.

CONCLUSIONS

Molecular and materials modeling is integrated into a substantial number of applications at DSM. Its integration into the project approach of R&D at DSM makes contribution easier through a developmental give-and-take using both quicker calculations and longer ones, both for quantitative and relative results. A strong interaction with the academic world is actively maintained, including project sponsorship and participation.

REFERENCES

Aagaard, O.J., R.J. Meier, and F. Buda. 1998. Ruthenium-catalysed olefin metathesis: A quantum molecular dynamics study. *J. Am. Chem. Soc.* 120, 7174-7182.

Barrett, R.A., and R.J. Meier. 1996. The calculation of molecular entropy using the semi-empirical AM1 method. *J. Mol. Struct.: Theochem.* 363, 03-209.

Coussens, B., K. Pierloot, and R.J. Meier. 1992. Conformations and rotational barriers of aromatic polyesters II, Rotational barriers in benzaldehyde and related aromatic moieties. *J. Mol. Struct.: Theochem.* 259, 331-344.

Hageman, J., R.A. de Groot, M. Heinemann, and R.J. Meier. 1997. The ultimate Young's modulus of orthorhombic polyethylene from *ab-initio* molecular dynamics. *Macromolecules* 30, 5953-5957.

Jahromi, S., B. Coussens, N. Meijerink, and A.W.M. Braam. 1998. Side-chain dendritic polymers: Synthesis and physical properties. *J. Am. Chem. Soc.* 120, 9753-9762.

Koglin, and R.J. Meier. 1999. An *ab-initio* study of the relative stability of the ggg and the gtg conformer in hexane. *Chem. Phys. Lett.* 312, 284-290.

Koglin, E., and R.J. Meier. 1999. Conformational dependence of vibrational frequencies and intensities in alkanes and polyethylene. *Computational and Theoretical Polymer Science.* 9, 327-333.

Meier, R.J. 1999. Applications of *ab-initio* molecular dynamics simulations in chemistry and polymer science. In *Computational Molecular Dynamics: Challenges, Methods, Ideas, Lecture Notes in Computational Science and Engineering* (M. Griebel, D.E. Keyes, R.M. Nieminen, D. Roose and T. Schlick, eds.) Springer-Verlag.

Meier, R.J. 1993. A comment on "Internal rotation in conjugated molecules: Substituted ethylenes and benzenes" by M. Head-Gordon and J.A. Pople. *Phys. Chem.* 97, 10248-10249.

Meier, R.J., and L.C.E. Struik. Atomistic modelling study of the relaxation processes in polymers: The beta-relaxation in polyvinylchloride. *Polymer* 39, 31-38.

Meier, R.J., A. Tiller, and S. Vanhommerig. 1995. Molecular modelling melamine-formaldehyde resins. 2. Vibrational spectra of methylolmelamines and bridged methylolmelamines. *J. Phys. Chem.* 99, 5457-5464.

Meier, R.J., G.H.J. van Doremaele, S. Iarlori, and F. Buda. 1994. *Ab-initio* molecular dynamics study of metallocene catalysed ethylene polymerisation. *J. Am. Chem. Soc.* 116, 7274-7281.

Reinhold, M., R.J. Meier, and C.G. de Koster. 1998. How feasible is MALDI-TOF-MS analysis of polyolefins? *Rapid Communications in Mass Spectrometry* 12, 1962-1966.

Sleigh, C., A.P. Pijpers, A. Jaspers, and R.J. Meier. 1996. On the determination of atomic charge via ESCA. *J. Electron Spectroscopy and Related Phenomena* 77, 41-57.

Tarazona, A., E. Koglin, F. Buda, B.B. Coussens, J. Renkema, J. van Heel, and R.J. Meier. 1997. Structure and stability of aluminum-alkyl cocatalysts in Ziegler-Natta catalysis. *J. Phys. Chem.* B 101, 4370-4378.

Terzis, A., D.N. Theodorou, and A.S. Stroeks. N.D.a. Entanglement network of polypropylene/polyamide interface 1. Network generation. *Macromolecules.* in press.

Terzis, A., D.N. Theodorou and A.S. Stroeks. N.D.b. Entanglement network of polypropylene/polyamide interface 1. Self-Consistent field model. *Macromolecules*, in press.

Wendling, J., A.A. Gusev, U.W. Suter, A. Braam, L. Leemans, R.J. Meier, J. Aerts, J.v.d. Heuvel and M. Hottenhuis. 1999. Crystal morphology and thermodynamics of poly(ethylene 4,4'-biphenyl dicarboxylate) and related copolymers with ethylene 2,6-naphthalene dicarboxylate. *Macromolecules* 32, 7866-7878.

Site:	**EniChem S.p.A.** **Centro Ricerche Novara "Istituto Guido Donegani"** **Via Fauser, 4; 28100 Novara** **Italy** **http://www.eni.it/english/panorama/organizzazione/enichem/**
Date Visited:	17 September 1999
WTEC Attendees:	K. Morokuma (report author), P.A. Kollman, E.B. Stechel, R. Miranda
Hosts:	Dr. Roberto Fusco, Coordinator, Molecular Modeling Department Phone: (39) 321-447325; Fax: (39) 321-447241 email: roberto.fusco@enichem.it Dr. Luca Longo, Molecular Modeling Department Dr. Paolo Accomazzi, Molecular Modeling Department Dr. Luigi Abbondanza, Molecular Modeling Department

BACKGROUND

EniChem is the chemical company of Eni Group (www.eni.it), which is owned 36% by the Italian government and 64% privately. Istituto Guido Donegani (IGD) is the corporate research center of EniChem.

It has approximately 250 employees, of which 100 are university-graduate researchers. The institute is organized under the director in a matrix structure, where four specialized departments (applied chemistry and physics, analytical chemistry and characterization, molecular modeling and process engineering) support the application projects carried out in five other departments (petrochemical intermediates, environmental technology, catalysis, physics and chemistry of polymers, and catalytic processes). The molecular modeling department has four research scientists, all of whom attended the site visit meeting.

R & D ACTIVITIES

Molecular modeling has a long history at IGD. A dedicated department was formed in 1974, when the research center was part of Montedison. At that time molecular modeling activity ranged from drug design to polymerization catalysis, involving about 20 people working under the supervision of Prof. E. Clementi. In 1979 the number of modelers was reduced to six, and Prof. C. Tosi became the leader of the department. In 1989 IGD passed from Montedison Group to EniMont (the joint venture between Montedison and EniChem) and then to EniChem. The business strategies of the new owner called for molecular modeling activities to focus mainly on polymer physics and polymerization catalysis. Until very recently, molecular modeling research was carried out in different departments. In 1999 an independent molecular modeling department was reconstituted, extending its activity to other kinds of catalysis, including those for organic intermediates from hydrocarbons.

The department leader, Dr. Fusco, used to work with Prof. Corradini of the University of Naples on modeling stereospecificity in olefin polymerization. Dr. Fusco joined IGD in 1982 for organic synthesis and later moved to molecular modeling. He works on polymerization catalysis, homogenous catalysis using molecular mechanics, conformational analysis and quantum mechanics. Of the other three members of the group, Dr. Abbondanza is involved in quantum mechanics, polymer crystal physics, statistical mechanics and software development. Dr. Accomazzi works on solid-state physics, polymer crystal physics, catalyst modeling and software development. Dr. Longo is involved in catalyst modeling, reaction mechanisms, molecular dynamics and meso-phase modeling.

The computer hardware at the institute is modest. There are two servers, with 4 and 2 cpu R10000, and 4 graphic workstations. As to the software, for quantum mechanics, they use DMol (MSI), ADF (SCM), Gaussian 98 , Gamess, Ampac and Mopac, Spartan. For the metallocene work described below, they used DMol, and more recently ADF and Gaussian. For molecular mechanics and molecular dynamics, they use

both home-grown codes and commercial packages like Cerius, Insight, Discover (all MSI), and Hyperchem. EniChem formerly was a member of the Polymer Consortium and later joined the Catalysis and Sorption Project, both of MSI. As a software development project, two members of the group are involved part-time in writing a quantum mechanics code, in which some nuclei (mainly protons) as well as electrons are handled quantum mechanically beyond the Born-Oppenheimer approximation.

Topics of research of the molecular modeling group have been (1) homogeneous catalysis, (2) olefin polymerization catalysis, and (3) polymer properties. They are also interested in heterogeneous catalysis.

Major recent results of the modeling group include the following:

- Metallocene catalyst activities. Mechanism of formation of active species, especially the role of cocatalysts.
- Correlation between catalyst structure and monomer reactivities in co- and ter-polymerization of ethylene, with other olefins and dienes. Some MO/MM hybrid methods have been used.
- Mechanisms of active species formation and deactivation in vanadium-based homogenous catalytic systems for olefin polymerization.

The most notable success stories for molecular modeling at EniChem come from metallocene catalysts. In one case, based on calculations, modelers proposed modified catalytic systems with improved performance. The experimental group tested the systems and they worked, resulting in two patents, with Fusco and Longo as primary authors. In the second case, an experimental group requested refinement of catalysts. Based on MM calculations, catalysts were improved and resulted in patents. The group members are convinced that close, and sometimes long-term, collaboration between modelers and experimentalists is essential to obtain good results.

DISCUSSION

These studies were undertaken in strong collaboration with the experimentalists. Some experimentalists are more open to theoretical suggestions than others. Theoretical studies have a good reputation within EniChem because of major accomplishments, and the goal is the design of new catalysts.

Good science and good application must still fit the business timing. For example, some years ago the molecular modeling group succeeded in developing a method, based on molecular dynamics simulations, to predict chain rigidity of liquid crystal polymers. However, by the time the research was completed, the company had lost interest in the subject. A side result of this study with interesting scientific implications was the observation that high-frequency vibrational modes of rigid polymers are highly unpopulated, and, consequently, the head-to-tail distance of polymer chains is shorter than the pure equilibrium statistics suggests (Fusco et al. 1993).

Publications are appreciated when the do not conflict with the need to protect the secrecy of results with strong industrial impact. The members of the modeling group have a rather large number of publications (see references as examples of most recent ones).

The need of the molecular modeling group at present is not to expand in size but to strengthen interaction with experimental groups. The collaboration has been extended to homogeneous and heterogeneous catalysts for production of intermediates starting with hydrocarbons.

Interaction with researchers outside the company is mainly via information from literature. There are some scientific collaborations with university scholars, most of whom are experimentalists. Interaction with university computational/modeling scholars has been rich in information content, and more such interaction is needed. New collaborations also are being considered for software development. For effective industrial/academic collaboration in general, university people should be more open to proposals from industry. The company is also participating in formal government-sponsored programs.

EniChem's modelers believe that a key need for molecular modeling software is a better transition-state search program. For instance, in some projects, they had to explore many side reactions, which required a better method. They are considering the use of quantum molecular dynamics methods to help to find alternative pathways in complex reaction systems.

CONCLUSIONS

Although the size is modest (four PhDs at present), the spirit of the molecular modeling group is very high. Through many years of collaboration with experimental groups, the modeling group has established itself as an indispensable component of the research center and has made many important contributions, including obtaining some patents, in particular in catalyst design.

EniChem's modeling efforts appear to have been fruitful both because of the expertise of its modeling staff members and because of their close interactions with experimental development. The group's successes have been significant, as recognized internally and externally. At Enichem, molecular modeling has reached the stage of being accepted as a normal, expected part of the technology innovation process in the company.

REFERENCES

Allegra, G., L. Caccianotti, and R. Fusco. 1998. Glass temperature of ethylene/propylene copolymers: An application of the pseudostereochemical equilibrium approach. *Macromolecules* 31, 5487-5494.

Fusco, R., and L. Longo. 1994. Thermodynamics of the early stages of homogenous Ziegler-Natta polymerization: density functional calculations on model systems $TiCl_2R^+$ and $TiCl_2R^+$. *Macromol. Theory Simul.* 3, 895-904.

Fusco, R., L. Longo, A. Proto, F. Masi, and F. Garbassi. 1998. Ethylene polymerization catalyzed by metallocene/methlaluminoxane systems: Quantum-mechanical study on the role of olefin separated ion pairs (OSIP) in the polymerization mechanism. *Macromol. Rapid Commun.* 19, 257-262.

Fusco, R., L. Longo, F. Masi and F. Garbassi. 1997a. Ethylene polymerization with homogeneous Ziegler-Natta catalysts: theoretical study on the role of ion pairs in the polymerization mechanism. *Macromol. Rapid Commun.* 18, 433-441.

Fusco, R., L. Longo, F. Masi, and F. Garbassi. 1997b. Olefin polymerization with homogeneous Ziegler-Natta catalysts: A DFT quantum-mechanical study of the reactions of Cp2MtCH3Cl complexes (Mt=Ti, Zr) with Al(CH3)3 and MAO. *Macromolecules* 30, 7673-7685.

Fusco, R., L. Longo, L. Caccianotti, C. Aratari, and G. Allegra. 1993. Molecular dynamics simulation of semirigid polymers. *Macromol. Chem. Theory Simul.* 2, 685-695.

Guerra, G., L. Cavallo, P. Corradini, and R. Fusco. 1997. Molecular mechanics and stereospecificity in Ziegler-Natta 1,2 and cis-1,4 polymerizations of conjugated dienes. *Macromolecules* 30, 677-684.

http://www.msi.com/solutions/cases/notes/enichem_full.html

Site: **EniTecnologie SpA**
Via F. Maritano 26
20097 San Donato Milanese
Italy
http://www.enitecnologie.it

Date Visited: 17 September 1999

WTEC Attendees: A. Tsao (report author), P.T. Cummings

Hosts: Francesco Frigerio, Department of Analytical Chemistry and Physical Chemistry 39 2 520 56443, 39 2 520 36347 Fax, ffrigerio@enitecnologie.eni.it
Dante Casati, Manager, Strategic Planning and Technology Innovation

BACKGROUND

EniTecnologie (formerly Eniricerche; http://www.enitecnologie.it/) is the corporate technology company within the EniGroup (Eni) with a mission of "generating value through technological innovation." Eni is the largest provider of energy products in Italy and a major provider in Europe and worldwide. Over the last five years, Eni has been undergoing restructuring following deregulation of the energy industry in Italy and government divestiture of 49% of its share of Eni.

EniTecnologie, in contrast to the research organizations within individual business units, is responsible for identification, investigation and technology transfer of research having high innovation and broad cross-cutting impact among business units within Eni. About 15-20% of the budget is directed toward long-range R&D. EniTecnologie consists of three technology centers: Catalysis and Process Technology, Chemistry and Product Technology, and Environmental Technologies, which are supported by two departments: Physical Chemistry and Engineering and Modeling. Molecular modeling is within Physical Chemistry.

MOLECULAR MODELING R&D ACTIVITIES

Molecular modeling at Eniricerche (now EniTecnologie) began in 1990. The first applications of molecular modeling were in protein engineering, materials modeling, and heterogeneous catalysis. At that time, biotechnology was an area of significant interest, which is no longer the case.

The budget is pre-allocated by the various business units and corporate with the confidence that commensurate value will be forthcoming. There are currently two researchers performing molecular modeling: Francesco Frigerio (full-time) and Roberto Millini (half-time), each of whom is involved in approximately two projects at any given time. There have always been two modelers. On projects in which they are involved, modelers are fully integrated members of a multidisciplinary research team, which can include theoreticians, experimentalists, and engineers. Projects are generally medium- to long-term in outlook. Modelers are readily acknowledged for their contributions on patents and among their customers within Eni. The contributions of modeling have created increasing demand for modeling, to the extent that demand exceeds the availability of modelers and resources.

Two long-term projects established the credibility of modeling within Eni through a sequence of consistent incremental achievements. One, in the area of protein engineering, sought to do NP-submodel building and modification of stability and activity. This project involved five people and lasted four years. A medium-term application to drug synthesis processes resulted in satisfying results for the pharmaceutical companies involved. In heterogeneous catalysis, contributions were made to a long-term project using molecular mechanics to study catalyst shape selectivity applied to the synthesis of cumene (Millini 1997), isobutene (Millini and Rossini 1997), and 2,6-dimethylnaphthalene by means of forced diffusion simulation and "binding energy" evaluation in zeolites. The modeling effort in this project helped identify interesting choices and resulted in a patent.

Short-term projects include applications to small-scale tests of innovative industrial processes (fast catalyst screening, pilot plants implemented). Other activities have involved solution of zeolite structure (Campbell et al. n.d.) and modeling of zeolite synthesis without (Millini et al. 1995) and with (Millini et al. 1998; Frigerio et al. 1998) templates.

Projects mainly use molecular simulation of solids and fluids (30%), molecular mechanics (MM) and dynamics (MD) (50%), and electron structure/MM hybrid methods (20%). Theoretically based correlations, quantum chemistry (QC) and solid-structure determination are occasionally applied. Mainly qualitative, rather than quantitative, understanding has been applied. QC is rarely used due to the cost of hardware and amount of human time that would be required. The majority of software used is purchased from commercial sources: Discover, Dmol, Gaussian, and MSI InsightII and Cerius2. One academic code (WHATIF) is also used. Limited-purpose software is developed in-house for simulation data analysis and presentation of results. Computational assets are limited to a collection of various types of SGI Unix workstations connected by the company LAN.

Near-term computational needs will be met with MM/MD and Monte Carlo, but it is anticipated that density functional theory and hybrid electronic structure/molecular simulation methods will be increasingly desirable in the long term. More accurate force fields and enhanced ability to model at the meso- and macroscale are highly desirable in the long term.

Although parallel computing would have a major impact, there is a critical need for codes that can be run on PCs and stand-alone workstations for both specialists and non-specialist use. Ease of use is currently of only moderate importance, but will be increasingly important in facilitating use by non-specialists.

Projects are often done in collaboration with other companies in the same business. EniTecnologie is also a member of the MSI Sorption and Catalysis Consortium. A small number of collaborations with academia have been attempted, and although a good model is still lacking, interest in such collaborations remains high.

CONCLUSIONS

There appears to be confidence that EniTecnologie will continue to support long-term research projects. In fact, it was felt that newly opening markets and environmental issues are resulting in an increased emphasis on innovation and long-term investment. The modeling group seems to be well established and has significant management advocacy. However, current staffing and resources would need to be increased if more demanding problems are to be addressed and if modeling were to be more extensively involved in the application field of EniTecnologie.

REFERENCES

Millini, R. 1997. The catalytic behavior of zeolites in cumene synthesis. *Solutions magazine*, MSI, January; http://www.msi.com/solutions/cases/notes/eniricerche_full.html.

Millini, R., and S. Rossini. 1997. In *Progress in Zeolite annd Microporous Materials.* H. Chon, S.-K. Ihm, U.S. Uh (eds.). 105. Elsevier Science Publishers. Amsterdam. 1389.

Campbell, B.J., A.K. Cheetham, D.E. Cox, R. Millini, G. Perego, L. Carluccio, and G. Bellussi. 1995. A new zeolite structure solved by simulated annealing. http://www.msi.com/solutions/cases/notes/anneal.html.

Millini, R., G. Perego, W.O. Parker, Jr., G. Bellussi, and L. Carluccio. 1995. *Microporous Materials* 4, 221.

Millini, R., Carluccio, F. Frigerio, W.O. Parker, and G. Bellussi. 1998. Zeolite synthesis in the presence of azonia-spiro compounds as structure-directing agents. *Microporous Materials* 24, 199.

Frigerio, F., L. Carluccio, R. Millini, and W.O. Parker. 1998. Molecular dynamics simulations of azoniaspiroalkanes in MOR and MTW. *Proceedings of the AIChE Topical Conference on Applying Molecular Modeling and Computational Chemistry*, Miami Beach, November 15-20, American Institute of Chemical Engineers, paper 52t.

Organization:	**GdR 12090 Project**
Meeting site:	TotalFina, European Technical and Research Center, Harfleur France
Date Visited:	15 September 1999
WTEC Attendees:	A.M. Chaka (report author), P.T. Cummings, K.K. Irikura, P. Vashishta, P.R. Westmoreland
Hosts:	Director of the GdR: Hervé Toulhoat (IFP) Senior-level members of the GdR and group leaders: François Hutschka (TotalFina), Jürgen Haffner (Vienna U.), Rutger van Santen (T.U. Eindhoven), S. Kasztelan (IFP), Daniel Bougeard (CNRS/UPRESA), Bernard Tavitian (IFP), János G. Ángyán (U. Henri Poincaré, Nancy), Christian Minot (U. Paris VI), Jean-Louis Rivail (U. Henri Poincaré, Nancy; CNRS representative) Others: Sylvain Clemendot (TotalFina), P. Raybaud (IFP), E. Payen (Lille) (Full addresses are given at the end of this report.)

BACKGROUND

Groupement de Recherche 12090 (GdR 12090) is a CNRS-sponsored consortium of industry, universities, and government labs whose purpose is to develop catalysis modeling tools for the petroleum industry. One primary goal is to create a community of scientists at the interface of simulation and heterogeneous catalysis. In addition to CNRS and French universities, members include IFP, TotalFina, T. U. Eindhoven, and the University of Vienna.

Fall 1999 is the end of the first four-year term of GdR 12090. The group was in the process of applying to the CNRS National Committee for a 4-year renewal for Phase II.

Research activities of individual members are briefly as follows:

- Haffner: LDA, VASP code development, applications to magnetic systems, non-periodic metals, catalysis
- Minot: Quantum chemistry, metal oxides/SiO_2 surfaces; collaborates with Toulhoat on chlorinated alumina; uses the program CRYSTAL
- Bougeard: Time-resolved spectroscopy; many computational methods; chemical reactivity; disordered materials; environmental chemistry
- Kasztelan: Applied work on hydrodesulphurization, hydrocarbon cracking, isomerization; mostly experimental, some modeling
- Tavitian: Thermodynamics, force fields for complex systems, zeolites, adsorption, vapor-liquid equilibria, viscosity; collaborates with Prof. Fuchs at U. Orsay; software developed at both IFP and Orsay
- Rivail: Solvent effects; enzymatic and heterogeneous catalysis (collaborating with Air Liquide)
- Ángyán: Theoretical chemistry, solvent effects, intermolecular interactions

DISCUSSION

The GdR was formed to more effectively transfer modeling technology from academia and national labs to industry. Flexibility and speed of interaction are seen as keys to its effectiveness. Many members present felt that MSI and commercial software in general are valuable links between academia and industry, and many use them. Unfortunately, commercial companies are often too slow to respond, and then users reach the limits of the code and cannot go beyond them. Universities respond faster and make the source code

available. Routine tools (i.e., commercial software) are best only for routine calculations. The expensive nature of MSI software is also a problem.

Initially, consortium members believed that Car-Parrinello quantum dynamics would be key, but that has proven to be a naive assumption. The calculations are still too time-consuming. However, developments in VASP, a plane-wave pseudopotential methodology, are cited as a key success of the GdR.

Jürgen Haffner commented that there are many benefits of industry working directly with universities. The universities receive funding and good research problems, and industry gets useful results. A serious problem, however, is the lack of permanent positions for the post docs and students trained in these fields. There are many postdoctoral opportunities in Europe in *ab initio* quantum mechanics and DFT. Unfortunately the opposite is true for permanent employment. Industry must be willing to hire, or universities will no longer be able to get students and postdocs and do this type of research.

To make the GdR more effective, facilitating transfer of modeling to industry in general, the government should provide funding to make academic-generated codes more user-friendly. Otherwise they will not be used.

CONCLUSION

This government-sponsored consortium, with its emphasis on industrial collaboration and applications, appears to be effective at developing useful molecular and materials modeling technology in academia and national labs and transferring it to industry. Software development is focused on solving specific types of problems of importance to industry. Participants are both modelers and experimentalists. There is a high degree of satisfaction among consortium members, as evidenced by the support for an additional four-year term for the program.

GDR 12090 PARTICIPANT ADDRESSES

F. Hutschka, TotalFina, Centre Europeen de Recherche & Technique, B.P. 27, 76700 Harfleur, France

C. Lamotte, J. Denayer, TotalFina, Research Center, Zone Industrielle C 7181 Seneffe (Feluy), Belgium

Hervé Toulhoat, P. Raybaud, Department of Thermodynamics and Molecular Modeling, Div. of Applied Chemistry and Physical Chemistry, Institut Français du Pétrole (IFP), 1 & 4 Avenue de Bois-Preau, 92852 Rueil-Malmaison, Cedex, France

S. Kasztelan, Director, Div. Kinetics & Catalysis, IFP (see above)

E. Payen, J.F. Paul, S. Cristol, Daniel Bougeard, Laboratoire de Catalyse Heterogene et Homogene, CNRS/UPRESA 8010, 59655 Villeneuve d'Ascq, France

G. Frapper, LACCO, Université de Poitiers, France

Jürgen Hafner, L. Benco, Institut fur Materialphysik and Center for Computational Materials Science, Universitat Wien, Sensengasse 8, 1090 Vienna, Austria

Jean-Louis Rivail, J. Ángyán, D. Parsons, Laboratoire de Chimie Theorique, UMR CNRS 7565 Institut Nanceien de Chimie Moleculaire, Université Henri Poincare, B.P. 239, 54506 Vandoeuvre-les-Nancy, France

Rutger van Santen, X. Rozanska, Eindhoven University of Technology, Department of Chemical Engineering, Laboratory of Catalysis, ST-W 4.28, 5600 MB Eindhoven, Netherlands

C. Minot, C. Martinsky, Markovits, Laboratoire de Chimie Theoretique, Université Pierre et Marie Curie, 4 place Jussieu, Boite 137, 75252 Paris, Cedex 06, France

M. Breysse, Laboratorie de Reactivité de Surface, Université Pierre et Marie Curie, 4 place Jussieu, Casier 178, 75252 Paris, Cedex 05, France

J. P. Gilson, M. Thibault, Mme Mauger, ISMRA—Laboratorie de Catalyse et Spectrochimie, 6 Bd. du Marechal Juin, 14050 Caen, France

Site:	**Glaxo Wellcome plc (now GlaxoSmithKline)** **Research and Development** **Medicines Research Center** **Gunnels Wood Road** **Stevenage, Hertfordshire SG1 2NY** **United Kingdom**
Date Visited:	September 13, 1999
WTEC Attendees:	P. Kollman (report author), R. Miranda, K. Morokuma, E. Stechel, A. Tsao
Hosts:	Dr. Michael M. Hann (mmh1203@glaxowellcome.co.uk) Andrew Leach Darren Green

BACKGROUND

Glaxo Wellcome plc (http://www.glaxowellcome.co.uk/) is one of the largest pharmaceutical companies in the world, with approximately $14 billion sales and $3 billion earnings in 1997. Approximately 50% of its worldwide research is done in Great Britain at the site visited. Subsequent to the site visit, it merged with SmithKline Beecham to become GlaxoSmithKline (http://www.gsk.com).

R&D ACTIVITIES

Mike Hann is the unit head for computational chemistry and informatics. Under him are three group leaders: Andrew Leach, head of the section on tool development methodology; Darren Green, head of the section on lead discovery; and Ian McClay, head of the section on lead optimization. The division of the modeling group into sections on lead discovery and lead optimization mirrors the division in the company (UK version) as a whole; other Glaxo Wellcome sites do not have the same organizational structure. There are currently 12 people in the modeling group, and six more are being sought to bring the number to 18. The expertise of the three scientists we met was diverse: Hann came from a background in synthetic organic chemistry, Leach had done conformational analysis, and Green was trained as a quantum chemist. The modeling group is funded out of central research.

The tools and methodology section headed by Leach concentrates on development of 3D databases, end-user tools for combinatorial chemistry, and analysis of reactivity in combinatorial reactions, as well as analysis of high-throughput, sometimes hard-to-quantify data for use in QSAR. The other two groups work closely with the project teams in lead discovery and optimization. Each person in the modeling group is, on average, on two project teams but also serves as consultant to others to help them use the modeling tools.

There are formal mechanisms to decide on target selection in defined disease areas; members of the computational chemistry group participate in this process. The number of targets with known structure has been increasing in the last years.

One area in which software has been developed internally is in combinatorial library design. This has been made possible by using the Web as a general interface to many processes (e.g., enumeration, profiling, selection) which are carried out by established computational chemistry methods. Often this has involved the use of commercial software packages (e.g., Daylight, MSI) being linked together through scripting languages.

Glaxo Wellcome has worked out a relationship with MSI to have someone on site to help integrate and develop MSI's software. This should in principle be a win-win situation because MSI software will have improved capabilities and Glaxo Wellcome will not have to keep integrating their developments again at each new code release. Glaxo Wellcome uses a large amount of commercial software, with about 30% MSI, 20% other companies, and the remaining home-built, with a large amount of the home-built using the

Daylight Toolkits. Glaxo Wellcome has been part of the Combinatorial Chemistry and Pharmaceutical Development consortia run by MSI.

The hardware is predominantly SGI, with a 32-processor Origin server for computational chemistry (just purchased at the time of the WTEC visit) and 3 Oracle servers on SGI, as well as approximately 80 SGI workstations (many deployed for bioinformatics) on site, which warrants the stationing of a SGI engineer on this site. There are also approximately 1000 PCs. The group would like to investigate the use of a "Beowulf"-like cluster; however, it appears that the computational chemistry efforts are not hardware-limited.

There are X-ray (5 people) and macromolecular NMR (2 people) groups located next to the computational group, so interactions are extensive and close. There are two combinatorial chemistry groups, one in lead discovery and one in lead optimization, whose efforts are supported by the computational chemistry group. There is some support for the product development group in an informal way and there is a separate group in bioinformatics, with 15-20 people, whose work overlaps a bit (e.g., homology modeling) with the computational chemistry group. The computational chemistry group provides consultancy to the four-person computational chemistry group in biometabolism. Thus, the computational chemistry group has lots of interactions, not only with chemists but also with other groups that support drug design efforts. There is also a chemoinformatics team being rebuilt within the IS division. This team will number about 5 people developing and supporting the use of more general informatics and data mining techniques for medicinal chemists

DISCUSSION

There is a wide range of views on the value of modeling among of the medicinal/synthetic organic chemists. Some view combinatorial chemistry as a rejection of rational design, but it is clear from the experience at Glaxo Wellcome that computational approaches have been very useful in combinatorial library design. The other extreme view, having computation totally drive design, is also rare, but the fact that there has been a mandate to significantly increase the size of the group is indicative of the appreciation of modeling as an important element in the design process. It was noted that one rarely considers a compound as coming exclusively from computational chemistry; computational modeling is more like an NMR machine, a crucial set of information that helps one proceed further and do a better job in design.

There is a "fairly open" attitude toward publication, and people are encouraged to do so; they can be rewarded for achieving international recognition. The computational chemists are considered for having their names on patents, with the process of who contributes to the inventive steps being considered. Whether or not they are listed on the patent, the important role of the computational group is acknowledged when a successful drug candidate emerges.

Glaxo Wellcome supports both in-house and external postdocs, supports CASE awards for graduate study, and encourages interactions with academia in both directions. The point was made that it is important for academics to be funded to stay in academia, because then their expertise in areas like statistics and genetic algorithms will flow into computational chemistry, improving the methodologies for all.

Some of the current areas where modeling could be very useful are the understanding of crystal forms, stability and formulation, and predicting metabolism, pharmacokinetics, solubility and homogeneity (in combinatorial chemistry). It would be very valuable to have a universal file format to allow more integration of the various software packages and to have more accuracy in docking and scoring, such as a real CASP-like prediction contest for ligand macromolecule interactions.

Site:	**Hoffmann-La Roche Ltd. (meeting site)** **Pharmaceutical Division; Chemical Tech-Mol. Design** **Grenzacher Strasse** **Basel, CH-4070, Switzerland**
Site:	**Novartis Pharmaceuticals AG** **Lichtstrasse 35** **Basel, CH-4002, Switzerland**
Date Visited:	16 September 1999
WTEC Attendees:	P.A. Kollman (report author), R. Miranda, K. Morokuma, E.B. Stechel, A. Tsao
Hosts:	Morning Session: Hoffman-La Roche Frank Grams (Frank.Grams@roche.com) Hans-Joachim Bohm Morning Session: Novartis Romain Wolf (romain.wolf@pharma.Novartis.com) Trevor Payne (trevor.payne@pharma.Novartis.com) Afternoon Session: Hoffman-La Roche Klaus Müller (klaus.mueller@roche.com) (with Kollman)

BACKGROUND

Hoffmann-La Roche (http://www.roche.com/) and Novartis (http://www.novartis.com/) are among the world's largest multinational pharmaceutical companies. Their headquarters are in Basel, but they both have very large research efforts in the United States as well. Novartis was formed by a merger of Ciba-Geigy and Sandoz.

R&D ACTIVITIES

Hoffman La Roche

At Hoffmann-La Roche (HLR), Frank Grams is the head of the molecular design group. It is made up of ten people, half doing modeling in support of projects (molecular modeling subgroup) and half focused on molecular property determination and analysis (molecular property subgroup). Dr. Grams was trained as a chemist and macromolecular crystallographer with Huber and was at Boehringer Mannheim, which was bought by HLR. He was brought in to head the molecular design group when Bohm was moved up.

Hans Joachim Bohm was head of the HLR molecular design group but now has moved to head the molecular structure unit, which includes small-molecule analytics, bioinformatics, chemical informatics, NMR, X-ray crystallography and modeling and design. Bohm was trained as a quantum chemist, was a modeler at BASF for 8 years, and has been at HLR 3 years.

The molecular design group is located in a separate building from medicinal chemistry, but participates as part of project teams, both in short- and long-term projects. There is quite a bit of program development by the team, most notably MOLOC, which emanated in the mid-1980s from an earlier modeling program by Klaus Müller and has been maintained and enhanced by Paul Gerber for about 15 years. This program is also used at other HLR sites and has been distributed to 12-15 academic sites. The five modelers in the molecular modeling subgroup work closely with chemists on specific projects; the other five (in the molecular property subgroup) are responsible for developing and enhancing the capabilities in physical property characterization of drug candidates and potential drug candidates.

The modeling group has a 16-processor SGI O2000, and each modeler has his own SGI O2 or Octane. The bioinformatics and crystallography groups also have their own SGI servers. In addition to their in-house

MOLOC code, they also use Macromodel, Insight, Daylight, Medchemxplorer, Xplor and Tsar and Corina from Oxford Molecular.

At HLR, 20-50% of targets have known structure, of which 10-50% are homology-model-built. Most of the HLR sites have active programs in modeling G-protein coupled receptors.

Worldwide, there are approximately 18 modelers in HLR. HLR also has macromolecular X-ray crystallography and NMR capabilities in-house. The modelers have strong interactions with the X-ray/NMR scientists.

HLR also has in-house combinatorial chemistry and collaborations with COMBICHEM and Arqule. Within the molecular design group, the group of five focusing on properties is also involved in library design/choice in order to make sure that what is chosen has suitable physical properties. HLR has bioinformatics co-located with modeling as a core technology, as well as collaborations with INCYTE and Affymetrix. HLR staff members see pharmacogenomics as a key technology of the future to target sub-populations with particular drug therapy that might not be suitable for everyone.

Novartis

At Novartis in Basel, each modeler is part of a disease area, with Romain Wolf and Trevor Payne part of the arthritis section. There are eight modelers at Novartis in Basel, four from the former Ciba-Geigy and four from the former Sandoz. They are also physically at two different locations. In contrast to Novartis Basel, the U.S. site of Novartis (like HLR) does not assign the modelers to specific project areas.

Wolf trained in biophysics and did modeling at Ciba-Geigy starting in the 1980s. Payne was trained as a medicinal chemist and is Chemical expert in the Arthritis group at Novartis. Wolf's primary duties are to support modeling in the arthritis area. Wolf stated that at first, he didn't think assignments by disease area would work well, but he has been pleased by how it has worked out. The arthritis area is characterized by its large number of targets whose structures are known, 90% or better. There is a large range in known structures, with 50% in the cardiovascular area and less in others. This percentage has clearly grown in the last few years. Like HLR, Novartis has macromolecular X-ray crystallography and NMR in-house, and modelers interact effectively with these other groups.

The modeling and bioinformatics groups have a 32-processor Origin 2000 and various SGI workstations. They have a wide variety of commercial software, and they spend on the order of $1 million annually worldwide.

DISCUSSION

Acceptance of modeling in these companies varies, and, as noted by Müller, there will be enthusiasts, people in the middle ground, and those not interested. He suggested that one must be able to give reasoned, but quick enough, answers to have an impact on the first two groups—a philosophy applied in the functionality of MOLOC. He recommended trying new modeling methodologies on the believers, emphasizing the importance of making the chemists understand that they share ownership with the modelers in the ideas that come from the modeling. There will always be nonbelievers, and one should not bother pursuing them. Bohm thought that the culture of HLR might be a little too individualistic and that the Roche prize could be given to teams rather than just individuals.

At HLR, credit is shared by project teams; however, patent law differs in different countries, so counting the modelers on patents may not be a good measure of "success" in their acceptance. HLR's thrombin patents include modelers, and there was one example mentioned at Novartis of a modeler on a patent.

At Novartis, publication is supported and encouraged. At a recent meeting at Novartis where prizes were awarded for the best posters, all the prize-winning posters showed NMR or X-ray structures of the targets.

As for outreach and contact with the outside world, Novartis has some collaboration with and support for PhD students, as well as collaboration with the University of Zurich. HLR is part of the Klebe consortium;

its researchers' attitude is that their chemical structures are crown jewels but that other things can be relatively openly shared. As noted, MOLOC is being used at universities. They have extensive contacts with ETH in Zurich. They provide some support for postdocs and summer students.

There was some skepticism about the value of consortia except in specific cases (e.g., the Klebe approach) where the companies get something back without just giving. For example, a consortium to focus on one metabolic path or on cytochrome P450 metabolism would be of interest.

Current and future challenges include the following:

- Reliable prediction of physical chemical properties (the focus of HLR's group of five within modeling)
- Better analysis of large and noisy data sets
- Improving docking and scoring to allow routine 10-fold enrichment over random screening (average now is 2 to 3, with 10 the best case)
- More experimental structural data for protein-ligand complexes
- Better ways to include the effects of water (solvation, hydrophobic effect)
- Improved force fields for modeling
- Structural concepts linked with automated chemistry
- Building up knowledge databases; for example, all metalloprotein binding sites or all pockets in proteins in terms of hydrogen bonding, properties, etc.
- Data mining in structural fashion for both chemists and biologists—e.g., all adenine binding sites, similarities and differences, and how they relate to the protein family

REFERENCES

Böhm, H.J., D.W. Banner, and L. Weber. 1999. Combinatorial docking and combinatorial chemistry: design of potent non-peptide thrombin inhibitor. *J. Comput.-Aided Mol. Design* 13, 51-56.

Böhm, H.J. 1998. Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3-D database search programs. *J. Comput.-Aided Mol. Design* 12, 309-323.

Stahl, M., and H.J. Böhm. 1998. Development of filter functions for protein-ligand docking. *J. Mol. Graph. Model.* 16, 121-32.

Stahl, M., D. Bur, and G. Schneider. 1998. Mapping proteinase active sites by projection of surface-derived correlation vectors. *J. Comp. Chem.* 20, 336-347.

Site: **IBM Zürich Research Laboratory**
Saynerstrasse 4
CH-8803 Reuschlikon
Switzerland
http://www.zurich.ibm.com

Date Visited: 17 September 1999

WTEC Attendees: E. Stechel (report author), P. Kollman, R. Miranda, K. Morokuma

Hosts: Dr. Wanda Andreoni (IBM), 41-1-724-8656, and@zurich.ibm.com
Prof. Ulrich W. Suter (ETH Zürich, Department of Materials), Tel: 41-1-632-3127, suter@ifp.mat.ethz.ch

BACKGROUND

International Business Machines Corporation, or IBM (http://www.ibm.com), was originally incorporated in the State of New York on June 15, 1911, as the Computing-Tabulating-Recording Co. (C-T-R), a consolidation of three existing companies. In 1924, C-T-R adopted the current name. In 1998, IBM had $81.7 billion in revenue, $6.3 billion in net income, $86.1 billion in total assets and a workforce of 300,000. It is one of the world's largest information technology companies.

IBM Research (http://www.research.ibm.com/) consists of eight research centers and laboratories worldwide organized around four themes: Smart Business, Innovative Interfaces, Deep Computing, and Serious Science:

- Smart Business
  - security
  - advanced commerce
  - mobile computing
  - network technologies
- Innovative Interfaces
  - multimedia
  - display technology
  - human-computer interaction
  - java
- Deep Computing
  - data storage
  - mathematics
  - data management
  - scalable computing
- Serious Science
  - chip technology
  - nanotechnology
  - computational biology
  - materials science

The three sites with "serious science" in their purview are the T.J. Watson and Almaden Research Centers, and Zürich Research Laboratory. Headquarters for the IBM Research Division is the T.J. Watson Research Center, in Yorktown Heights, N.Y. It was established in 1961 and houses ~1,400 employees with a focus on

semiconductors, physical and computer sciences, and mathematics. The Almaden Research Center in San Jose, California was established in 1986 and has ~700 employees with a focus on storage systems and technology, computer science, and science and technology.

Of the remaining five sites, two were established in 1982:

- The Haifa Research Laboratory, Haifa, Israel, has ~300 employees who focus on systems technology, verification technologies, e-business and security, system applications, languages and optimization, applied mathematics, multimedia, and VLSI design technologies.
- Tokyo Research Laboratory in Yamato, Japan has ~175 employees who focus on software technology, system technology, network and solutions technology, and computer science theory.

The three remaining laboratories were established recently and are still very small:

- Austin Research Laboratory in Austin, Texas, established in 1995, has ~30 employees.
- China Research Laboratory in Beijing, China, also established in 1995, has ~60 employees.
- Delhi Solutions Research Center in Delhi, India, was established in 1997 and has ~20 employees.

The Zürich Research Laboratory in Rueschlikon, Switzerland is the oldest of the eight sites; it was established in 1956. It has ~200 employees who focus on communication systems and related technology solutions, optoelectronics and physical sciences. The organization of the laboratory is divided into three areas:

1. Communication systems: In communications research, the laboratory focuses on the field of asynchronous transfer mode (ATM), which is the basic technology for emerging high-speed networks, with multimedia capabilities in local as well as wide area environments. A second focus is in wireless networking—i.e., transmission systems for data communication using radio frequencies and infrared.
2. Information technology solutions: The focus here is to develop solutions and applications and provide services to customers in networking-related fields.
3. Science and technology: Current projects include micro- and nanomechanics, display technology, microcontact processing, optical networking technology, transport and magnetism in layered structures, nanoscale science, and computational materials science.

R&D MODELING ACTIVITIES

Computational molecular-scale modeling at IBM Zürich currently consists of three people, Wanda Andreoni (group leader), Peter Blöchl, and Alessandro Curioni. At the time of the WTEC visit they were actively seeking to fill two open positions. They also augment most of their efforts through active collaborations. For example, our co-host, Ulrich Suter, a polymer chemist from the Department of Materials at Eidgenössische Technische Hochschule (ETH) Zurich, is a close collaborator. The group's name is "computational materials science" and it is a part of the science and technology department, which has a total of ~38 people. The name has changed through the years, from "physics & chemistry" to "chemistry & physics" to the current name. The name changes have reflected changes in research focus. The most important recent change is a switch to a problem focus, away from a focus on methods.

Management and staff members intend for the computational modeling effort to be state-of-the-art, respected by the scientific community as technically excellent. Management also expects that they will transfer to customers the Car-Parrinello Molecular Dynamics technology (both software, through licensing, and the SP computer architecture that it is optimized for). In particular, their research is expected to provide unique demonstrations of how CPMD and their expertise in computational physics and chemistry can be used for practical applications, including engineering of new materials and chemical processes.

Researchers are free to choose their problems and collaborators, consistent with promoting high-performance computing or materials research of internal relevance; and they are evaluated based on the quality of their

scientific contributions. Personal expectations include intellectual challenge, variety, excitement, and satisfying scientific output.

Molecular modeling projects at IBM Zürich include the following:

- Organic and metal-based drugs interacting with nucleobases and oligonucleotides
- Adsorption and chemical reactions in zeolites
- Chemistry and dynamics of fullerene derivatives
- Studies in catalytic activity
- Role of lithium in the polymerization of synthetic rubber
- Transition metal organometallics that catalyze the synthesis of chiral compounds
- Light-emitting organic materials for display applications

Work is done primarily on the SP architecture and they are awaiting the arrival of the SP3. IBM modelers use CPMD 98% of the time. They have Gaussian 98, mostly for testing small molecules, and MOPAC for comparisons. They also have Cerius and AM1. They strongly oppose any black-box software packages.

Suter's Polymer Chemistry Group at ETH Zürich focuses on modeling polymers from the atomistic level to the continuum, polymers near interfaces, and functional polycondensates.

DISCUSSION

IBM researchers have a large number of active and intense collaborations. Indeed, this is simply the way they do research—i.e., their mode of operation. For example, when Michele Parrinello first came to IBM Zürich, he visited and interviewed professors in every University in Switzerland as potential collaborators. Since Parrinello's departure, the group has maintained this culture. They have generally seen very little value in consortia or virtual networks. However, they are a member of the European Science Foundation's SIMU network of molecular simulations groups (http://www.esf.org/physical/pp/SIMU/) and are the contact point in Switzerland.

Breakthroughs that could have large impact include the following:

- Combining methodologies
- Scale-marrying approaches
- Methodologies for information harvesting from large computational data sets
- "Marrying" chemical reactions and the physics of materials

CONCLUSIONS

Because IBM sells computers and computer solutions, the scientific community and other industries are all potential customers for this research group. This situation gives IBM researchers a rich variety of potential research projects and many opportunities to work on state-of-the-art technologies and have impact on their company.

REFERENCES

Blöchl, P.E., and J.H. Stathis. 1999. Hydrogen electrochemistry and stress-induced leakage current in silica. *Phys. Rev. Lett.* 83, 372.

Curioni, A., and W. Andreoni. 1999. Metal-Alq3 complexes: The nature of the chemical bonding. *J. Am. Chem. Soc.* 121, 8216.

Curioni, A., W. Andreoni, R. Treusch, F.J. Himpsel, E. Haskal, P. Seidler, C. Heske, S. Kakar, T. van Buuren, and L.J. Terminello. 1998. Atom-resolved electronic spectra of Alq3 from theory and experiment. *Appl. Phys. Lett.* 72, 1575.

Curioni, A., M. Boero, and W. Andreoni. 1998. Alq3: *ab initio* calculations of its structural and electronic properties in neutral and charged states. *Chem. Phys. Lett.* 294, 263.

Haskal, E.I., A. Curioni, P.F. Seidler, and W. Andreoni. 1997. Lithium-aluminum contacts for organic light-emitting devices. *Appl. Phys. Lett.* 71, 1151 (1997).

Hass, K.C., W.F. Schneider, A. Curioni, and W. Andreoni. 1998. The chemistry of water on alumina surfaces: Reaction dynamics from first principles. *Science* 282, 265.

Senn, H.M., P.E. Blöchl, and A. Togni. 1999. Towards an alkene hydroamination catalyst: static and dynamic *ab initio* DFT studies. Submitted to *J. Am. Chem. Soc.*

Site: **Institut Français du Pétrole (IFP)**
1 & 4 Avenue de Bois-Preau
92852 RUEIL-MALMAISON Cedex FRANCE
http://www.ifp.fr
(Joint visit with GdR 12090 Project at TotalFina Raffinage)

Date Visited: 15 September 1999

WTEC Attendees: A.M. Chaka (report author), P.T. Cummings, K. Irikura, P. Vashishta, P.R. Westmoreland

Hosts: Dr. Hervé Toulhoat, Group Leader, Molecular Modeling and Computational Chemistry, herve.toulhoat@ifp.fr
Bernard Tavitian (Dept. of Thermodynamics), tavitian@ifp.fr, 1 & 4, avenue de Bois-Preau; 92852 Rueil-Malmaison, Cedex, France; Phone: 33 1 47 52 60 00; FAX: 33 1 47 52 70 00

BACKGROUND

The IFP (http://www.ifp.fr) is a primarily tax-supported institute dedicated to research and development, training and education, and information management for the oil, gas, and automotive industries. It is also the leading independent developer of automotive engines in France. The IFP currently has partnerships with approximately 30 oil and gas companies, 40 supply and service firms, and 20 automotive manufacturers and component suppliers.

Statistics from the 1998 annual report are as follows:

- 1,820 employees
- Budget of 1.9 billion French francs (291 million Euro)
- 1,301 patents filed
- 14,000 total "living" patents
- 124 industrial development contracts (93% signed abroad)/110 ongoing theses
- 362 students enrolled, 45% from other countries
- 30 post-docs annually
- 827 continuing education courses
- 328 scientific articles and papers published
- 17 books published

Technology is transferred to industry in two principal ways. One is through the licensing of technology to established companies. By 1998, IFP had negotiated and signed over 900 licenses with industry. The other is through starting small companies to develop and market the technology. IFP retains 52% of the stock in these startup companies, formally known as the ISIS group.

Research and development is structured around four fundamental areas of interest to the oil and gas industry: exploration-reservoir engineering, drilling-production, refining-petrochemicals, and internal combustion engines and energy. This research is conducted within 12 research divisions.

There is a high value placed on advanced fundamental and exploratory research. The focus of the exploratory research is in five areas: catalysis, separation, analysis, process engineering, and materials. Research has a centralized structure and emphasizes a multidisciplinary approach to solving problems. There is a large computational fluid dynamics research effort focused on exploration and production, formation and migration of hydrocarbons.

Environmental protection is a major driver of research programs at IFP. For the refining segment a key focus of research is reducing the level of sulfur in both gasoline and diesel fuels to reduce SO_2 emissions. Hence, dehydrosulfurization catalysts are very important.

MODELING R&D ACTIVITIES

Molecular modeling at IFP began in 1988. History and highlights may be summarized as follows:

- 1988: First workstation dedicated to molecular modeling (Chem-X on a VAX station)
- 1989: Beginning of participation in three Biosym consortia: Potential Energy, Catalysis & Sorption, and Polymers
- 1990: Two SGI 4D35 workstations purchased
- 1991: Creation of a specific project within the prospective research program
- 1993: Joined the Biosym EOM consortium
- 1993-1994: H. Toulhoat at Total
- 1995: Formal creation of the molecular modeling group within DISMA (Direction des l'Informatique Scientifique et des Mathématiques Appliquées)
- 1995: "Molecular Modeling in Oil and Gas Sciences" organized by the IFP
- 1996: Launch of the GdR 1206 (CNRS, IFP, TUW)
- 1998: Total rejoins the GdR 1206

The molecular modeling group in DISMA is led by Dr. Hervé Toulhoat (Thèse ENSM Paris 1980) and includes Dr. Pascal Raybaud (UPMC/TUW 1998) and Dr. Anne-Dominique Bianciotto-Defontaines (UPMC 1992). It also includes a postdoc (Dr. Magdalena Dymitrowska, U Paris XI Orsay 1998), two thesis students, and three trainees.

Others at IFP also carry out molecular modeling:

- The Department of Thermodynamics (Division of Applied Physical Chemistry and Materials) has 2 engineers doing modeling more than 50% of their time.
- The Department of Geochemistry (Geology Division) has 2 engineers doing modeling part-time. The Physical and Analytical Division has no assigned specialists but is involved in selected applications.
- The Kinetics and Catalysis Division has numerous collaborations with the DISMA modeling group and is involved in dissemination of desktop tools.

IFP employs a broad range of molecular and materials modeling tools, from empirical force fields for polymer simulations, to Car-Parrinello quantum dynamics for heterogeneous catalysis, to quantitative structure-property relationships (QSPR). A wide range of hardware and software is available, as shown in Table B1.6.

Except for a few very specific needs, new molecular modeling methodology is not developed at IFP because of the vast array of methods currently available commercially or from universities. IFP relies on partnerships to do development work, such as MSI consortia or the GdR 12090 project. A project to develop Order-N scaling for quantum mechanics is conducted in conjunction with the IRC, University of Madrid, CEA, EDF, and CNRS. In addition, IFP has collaborated with LPCMA (Prof. Fuchs, Orsay) since 1994, and sponsors PhD CIFRE theses with INRIA.

Some development work is being done to obtain better force field parameters for multipolar molecules (sulfur-containing, aromatics, etc.) and cycloalkanes, as well as new algorithms for the direct determination of bubble points. Particularly challenging systems for modeling VLE (vapor-liquid equilibrium) and transport properties are gas condensates and long-chain molecules such as squalenes.

Table B1.6
IFP Hardware and Software Resources

Hardware	Shared resources: NEC SX-5 (5 procs, 40 GFlops, 30 Go) SGI-CRAY Origin 2000 (20 procs, 8 GFlops, 10 Go) High-speed network
	Divisional resources: DISMA molecular modeling group: 6 SGI R10000s RG Division: 2 SGI R10000s RB Division: 1 SGI R10000
Graphical Interfaces	InsightII 400P (MSI) Cerius2 4.0 (MSI)
Quantum Mechanics Codes	VASP (DFT, planewave pseudopotential) (Univ. of Vienna) SIESTA (DFT, Order N, pseudopotential) (Univ. of Madrid) DMol3 (DFT, numerical basis set) (MSI) CASTEP (DFT, planewave pseudopotential) (MSI) MOPAC 6 (MSI interface) Gaussian 98 (source code) (Gaussian Inc.) Various codes from QCPE
Molecular Mechanics Codes	Discover, OFF (MD) (MSI) DL_Poly (MD) (Daresbury) Sorption, Flexisorb (GCMC) (MSI) Gibbs (GEMC) (LPCMA)
Mesoscale Software	MESODYNE and DPD (MSI)
Programs that Calculate Spectra	EXAFS (Daresbury/MSI under Cerius2) TEM (MSI/Cerius2) IR/RAMAN (MSI/Cerius2) X-ray Diffraction (MSI/Cerius2 and InsightII) NMR (PolyNMR/InsightII, *ab initio*) UV/VIS (G98, ESOCS for solids/MSI)
Programs to Predict Properties	QSPR (MSI/Cerius2) Polymers: InsightII 400P, Synthia, Van Krevelen

Modeling applications and achievements are summarized in Tables B1.7 and B1.8, respectively, and in the accompanying references.

Table B1.7
Applications of Molecular and Materials Modeling at IFP

Thermodynamics	Calculation of phase diagrams (liquid/vapor, liquid/solid, heavy products etc.) Adsorption isotherms for zeolite separations Calculating enthalpies and entropies for development of new process
Kinetics	Evaluating reaction mechanisms and characterizing transition states Calculating activation energies and frequency factors for process modeling Calculating the effects of isotopes for petroleum geochemistry
Catalysis (heterogeneous and homogeneous)	Chemistry of synthetic processes Models of active phases and active sites Relationship between adsorption and the dimension of micropores Study of elementary reaction mechanisms Analysis of steric constraints, e.g., reactions in zeolite cages Poisoning
Materials	Reactivity/corrosion (metals) Adhesion and bonding (polymers, metals) Mechanical properties (polymers) Molecular and colloidal additives Additives for fuels and lubricants Additives for exploration (formulation of boues) Additives for production (e.g., hydrates, paraffins) Processes for assisted recovery (polymers and TA, aqueous)
Hydrocarbon Chemistry	Thermal cracking (generation and maturation of petroleum, refining and petrochemicals) Combustion (boilers, engines, pollution minimization) Modeling of heavy fractions (asphaltenes, resins, *kérogèn*) Interactions at interfaces (migration, retension, evaluation of reserves) Molecular identification (calculating IR, UV/VIS, NMR spectra by *ab initio*)

Table B1.8
Selected Accomplishments

Production	Dehydrosulfurization catalysts: active sites, effect of promoter, energy profiles of reactions (Raybaud et al. 1997a; Raybaud et al. 1997b; Raybaud 1998; Raybaud et al. 1998a; Raybaud et al. 1998b; Toulhoat et al. 1999; Raybaud et al. 1999a; Raybaud et al. 1999b) Mechanism of sulfur poisoning and chlorine restoration of the activity of palladium hydrogenation catalysis (Valerio and Toulhoat 1996; Valerio and Toulhoat 1997; Gravil and Toulhoat 1999a; Gravil and Toulhoat 1999b) Hydrocracking (thesis in progress) Origin of selectivity in distillation media Morphology of precipitated Boehmite (Wellens 1997; Raybaud et al. 2001) Nanoparticle morphology as a function of T, pH, and additives from measures from *ab initio* molecular dynamics simulation of the solid/solution interfacial tension Separation mechanism of o-xylene/p-xylene in zeolites X (Mackie et al. 1997; Lachet et al. 1998a; Lachet et al. 1998b; Lachet et al. 1998c; Lachet et al. 1999) Mechanical properties of crystalline and amorphous polyamides (Dal Maso et al. 1997) Atomistic calculation of elastic constants Permeability of PVDF to CH_4 and CO_2 (flexible pipes) (Bianciotto-Defontaines and Toulhoat 1998a; Bianciotto-Defontaines and Toulhoat 1998b) Molecular mechanisms, atomistic calculations of S and D (P=S*D) Kinetic inhibitors of gas hydrates (Bianciotto-Defontaines et al. 1999)
Exploration	Hydrocarbon liquid/vapor equilibrium (Neubauer et al. 1999a; Neubauer et al. 1999b; Neubauer et al. 1999c; Ungerer 1999) Reproducing phase diagrams, force fields AUA Molecular modeling of asphaltenes (Toulhoat 1996; Jeuland 1997)

(Note that the open discussion from this site visit is included in the discussion section of the GdR 12090 report.)

REFERENCES

Bianciotto-Defontaines, A.D., and H. Toulhoat. 1998a. Evaluation des coefficients de diffusion de CH4 et CO2 dans le PVDF avec les outils de la modélisation moléculaire. *Rapport IFP* N° 45146 (December).

Bianciotto-Defontaines, A.D., and H. Toulhoat. 1998b. Evaluation de la solubilité de CH4 et CO2 dans le PVDF avec les outils de la modélisation moléculaire. *Rapport IFP* N° 45021 (October).

Bianciotto-Defontaines, A.D., H. Toulhoat, and E. Dehaudt. 1999. Etude par modélisation moléculaire de l'adsorption d'inhibiteurs de croissance cristalline à l'interface eau/hydrate. *Rapport IFP* N° 45505 (May).

Dal Maso F., J. Benkemoun, and H. Toulhoat. 1997. Possibilités actuelles du calcul des constantes élastiques de polymères par des méthodes de simulation atomistiques. *Revue de l'IFP* 52, 6, 625-641 (1997).

Gravil, P., and H. Toulhoat. 1999a. Hydrogen, sulfur and chlorine co-adsorption on Pd(111): a theoretical study of poisoning and promotion. *Surface Science* 430, 176-191 (1999).

Gravil, P., and H. Toulhoat. 1999b. Ethylene, sulfur and chlorine co-adsorption on Pd(111): A theoretical study of poisoning and promotion. *Surface Science* 430, 192-198.

Iftimie, R., P. Raybaud, and H. Toulhoat. 1998. Dynamique moléculaire à l'interface boehmite-eau. *Rapport IFP* N°44748 (July).

Jeuland, N. 1997. Modélisation moléculaire d'additifs détergents pour carburants: solvatation et adsorption à l'interface paraxylène/noirs de carbone de dérivés du polyisobutène-succinimide. *Rapport de Stage Élève-Ingénieur ENSCP* (July).

Lachet, V., A. Boutin, B. Tavitian, and A.H. Fuchs. 1997. Grand canonical Monte Carlo simulations of adsorption of mixtures of xylene molecules in faujasite zeolites. *Faraday Discuss.* 106, 307-323.

Lachet, V., A. Boutin, B. Tavitian, and A.H. Fuchs. 1998a. Adsorptive separation of aromatics in faujasite zeolites: A grand canonical Monte Carlo simulation study. Preprints of the topical conference on separation science and technologies, part II, Am. Institute of Chem. Eng., New York, 1080-1085.

Lachet, V., A. Boutin, B. Tavitian, and A.H. Fuchs. 1998b. Computational study of p-xylene/m-xylene mixtures adsorbed in NaY zeolite. *J. Phys. Chem.* B, 102, 9224-9233.

Lachet, V., S. Buttefey, A. Boutin, B. Tavitian, and A.H. Fuchs. 1998c. Molecular simulation of xylene isomers adsorption in faujasite zeolites. *Proceedings Of The 6th International Conference On Fundamentals of Adsorption*, Francis Meunier (Ed.), 879-884. Elsevier.

Lachet, V., A. Boutin, B. Tavitian, and A.H. Fuchs. 1999. Molecular simulation of p-xylene and m-xylene adsorption in Y zeolites. *Langmuir* in press.

Mackie, A.D., B. Tavitian, A. Boutin, and A.H. Fuchs. 1997. Vapour-liquid phase equilibria predictions of methane-alkane mixtures by Monte Carlo simulation. *Molecular Simulation* 19, 1-15.

Neubauer, B., J. Delhommelle, A. Boutin, B. Tavitian, and A.H. Fuchs. 1999a. Monte Carlo simulations of squalane in the Gibbs ensemble. *Fluid Phase Equilibria* in press.

Neubauer, B., B. Tavitian, A. Boutin, and P. Ungerer. 1999b. Molecular simulations on volumetric properties of natural gas. *Fluid Phase Equilibria* 161, 45.

Neubauer, B., A. Boutin, B. Tavitian, and A.H. Fuchs. 1999c. Gibbs ensemble simulations of vapour-liquid equilibria of cyclic alkanes. Submitted to *Molecular Physics*.

Raybaud P., G. Kresse, J. Hafner, and H. Toulhoat. 1998a. The adsorption of thiophene on the catalytically active edge-surface of MoS2: An *ab-initio* local-density-functional study. *Physical Review Letters* 80, 7, 1481-1484.

Raybaud P., J. Hafner, G. Kresse, and H. Toulhoat. 1998b. Structural and electronic properties of the MoS2 (10-10) edge-surface. *Surface Science* 407, 237-250.

Raybaud P., G. Kresse, J. Hafner, and H. Toulhoat. 1997a. *Ab-initio* density-functional studies of transition-metal sulfides: II. Electronic structure. 1997b. *Journal of Physics, Condensed Matter* 9, 11107-11140 (1997).

Raybaud P., G. Kresse, J. Hafner J., and H. Toulhoat. 1997b. *Ab-initio* density-functional studies of transition-metal sulfides: I. Crystal structure and cohesive properties. *Journal of Physics Condensed Matter* 9, 11085-11106.

Raybaud, P., and H. Toulhoat. 1998. Estimation par modélisation moléculaire des enthalpies de formation des constituants du modèle de reformage catalytique. *Rapport IFP* N° 45137 (December).

Raybaud, P., J. Hafner, G. Kresse, S. Kasztelan, and H. Toulhoat. 1999a. *Ab-initio* study of the H2-H2S / MoS2 gas-solid interface: the nature of the catalytically active sites. *J. Catalysis* (submitted, June).

Raybaud, P., J. Hafner, G. Kresse, S. Kasztelan, and H. Toulhoat. 1999b. *Ab-initio* study of the structural, energetic, and electronic properties of the edge surfaces of promoted MoS2 catalysts. *J. Catalysis* (submitted, July).

Raybaud, P., M. Digne, R. Iftimie, W. Wellens, P. Euzen, and H. Toulhoat. 2001. Morphology and surface properties of boehmite (•-AlOOH): A density functional theory study. Accepted, *J. Catal.*

Toulhoat, H. (Coordinateur). 1996. La modelisation moléculaire dans les sciences pétrolières et gazières. Numéro spécial: actes des rencontres scientifiques de l'IFP, *IRevue de l'IFP* 51, 1.

Toulhoat, H., P. Raybaud, S. Kasztelan, G. Kresse, and J. Hafner. 1999. Transition metals to sulfur binding energies relationship to catalytic activities in HDS back to Sabatier with first principles calculations. *Catalysis Today* 50, 3-4, 629-636.

Ungerer, P., C. Beauvais, J. Delhommelle, A. Boutin, B. Rousseau, and A.H. Fuchs. 1999. Optimization of the anisotropic united atoms intermolecular potential for n-alkanes. Submitted to *J. Chem. Phys.*

Valerio, G., and H. Toulhoat. 1997. Atomic sulfur and chlorine interaction with Pdn clusters (n=1-6): a density functional study. *J. Phys. Chem.* A, 101, 1969-1974.

Valerio, G., and H. Toulhoat. 1996. Local, gradient-corrected, and hybrid density functional calculations on Pdn clusters for n=1-6. *J. Phys. Chem.* 100, 26, 10827-10830.

Wellens, W. 1997. Modélisation moléculaire de propriétés de surface d'alumines utilisées en catalyse hétérogène. *Mémoire de DEA Cinétique, Catalyse, Combustion et Modélisation ENSPM/UPMC* (June).

Fabienne Chevereau, F. 1997. Les asphaltènes en solution: recherche de modèles moléculaires, et relations entre structure et solubilité. *Mémoire de Fin d'Études d'Ingénieur ISMANS, Option CAO Moléculaire* (September).

Sites:	**The Netherlands, including TU Eindhoven, Royal Dutch Shell, Philips**
Visit Location:	Workshop at Eindhoven University of Technology (TU Eindhoven)
Date Visited:	17 September 1999
WTEC Attendees:	A.M. Chaka (report author), K.K. Irikura, P. Vashishta, P.R. Westmoreland
Participants:	Prof. Rutger Van Santen, Eindhoven University of Technology, Laboratory of Inorganic Chemistry and Catalysis P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands Phone: 31 40 247 2730; Fax: +31 40 245 5054 R.A.v.Santen@tue.nl Dr. Anik Peeters, University of Antwerp, anik@uia.ac.be

BACKGROUND

This report on activities in the Netherlands is organized differently than the other site reports because it is derived from the literature, the Web, and discussions at a workshop at the Eindhoven University of Technology (*Technische Universiteit Eindhoven*, TUE). The workshop covered the breadth of work in the Netherlands and in Belgium (the latter reported separately). This report also includes general remarks on modeling of heterogeneous catalysts.

Much of the discussion of molecular modeling in the Netherlands focused on industrial use. In the Netherlands, molecular modeling is now making a comeback in academia, but Shell and Philips no longer appear to be conducting work in the area. Philips presently has about 9000 people in its research labs, and yet has no molecular modeling activity. Philips had continued molecular modeling until quite recently, including work with Car-Parrinello molecular dynamics. Unilever at Vlaardingen has work on melting by Agterof and more general molecular modeling activities by de Groot. DSM has an active program, described in a separate site report.

Attendance at the workshop at Eindhoven was small, as several other conferences were occurring at the same time. In addition, Prof. van Santen felt that the small attendance was also due to the fact that most of the academics in the Netherlands are focused on increasing the accuracy of theoretical methods and are uncomfortable with the approximations often necessary to solve industrial problems. They are mostly interested in the pure, not the applied, science. ADF, the Amsterdam Density Functional Program from Behrends' group, is the biggest exception. Rhône-Poulenc has expressed the same frustration in trying to collaborate with French academics.

CATALYSIS MODELING R&D AT TU EINDHOVEN

Prof. Dr. Rutger van Santen hosted the workshop at Eindhoven. Because of his range of experience, Prof. van Santen is well positioned to know much of the Dutch academic and industrial modeling community. He received his PhD from Leiden University in 1971 as a theoretical organic chemist with Prof. L.J. Oosterhoff and then was a postdoctoral fellow in 1971-72 at SRI (Menlo, California) in the Molecular Physics Department. Next he joined the Koninklijke/Shell Laboratorium Amsterdam, where he had several job functions in catalysis research. During 1982-84, he was assigned to Shell Development Company in Houston, Texas. In 1986 he joined the Chemical Engineering Department at Eindhoven University of Technology as a part-time professor in surface chemistry, moving to full-time status in 1988 as professor in catalysis.

Professor van Santen's current research interests are in the molecular aspects of heterogeneous catalysis with three main themes: computational studies of surface-chemical reactivity, mechanisms in heterogeneous catalysis, and physical chemistry of catalyst synthesis. In his lab, there are eight PhD students and two post-doctoral associates working in theory, half in quantum mechanics and half in kinetic Monte Carlo

simulations. His group uses the quantum-mechanics codes VASP to get geometry and ADF to do the analysis of the electronic structure and to develop chemical interpretation of bonding.

Prof. van Santen receives a large amount of his funding from industry and spends approximately half his time interacting with industry through company sponsorship of students. He currently has contracts with companies such as Solvay, DuPont, Norsk Hydro, and TotalFina. He is also a participant in the GdR 12090 project, described in a separate site report.

His laboratory is part of the multi-department Schuit Catalysis Institute founded in 1960 (SCI, http://www.tak.chem.tue.nl/ski.html), itself a participant in the Netherlands Institute of Catalysis Research (NIOK, http://www.chem.tue.nl/niok/). In the SCI, there are five full professors in reaction engineering, one in information science, and two in technical physics. In the Dutch academic system, the groups are very large, and a full professor has on average 2-3 faculty members around him. Thus in SCI, there are 30 faculty and a total of 60 students. NIOK is on an even larger scale, linking graduate research on catalysis at the Universities of Amsterdam, Delft, Eindhoven, Leiden, Groningen, Utrecht, and Twente.

An interesting aspect of Dutch academia is that a substantial amount of direct funding may come from industry. Out of the 70 people in the chemical engineering department, six lecturers and three professors are funded by industry. In the 1970s, half of the faculty members at Dutch universities had relationships with Shell, Philips, or Unilever in that they spent 2-3 days a week at the companies as employees. Equipment was much better at these companies than in the universities.

ROYAL DUTCH SHELL

Professor van Santen offered his observations and thoughts about the history of molecular modeling at Shell. In the 1960s, managers at Shell and Philips believed in quantum chemistry. There was quantitative failure in quantum mechanics at that time, however. Extended Hückel worked fairly well for organic systems, but the NDO (neglect of differential overlap) semi-empirical approximation did not provide good enough results. As a result, managers from that period came to believe that molecularly based modeling doesn't work.

When van Santen came back to Royal Dutch Shell in 1984 from Houston, there was a revolution in modeling going on. In the 1980s, an individual whose background was the mathematics of valence bond theory became the new director of research at Shell. He saw supercomputers being used by oil companies for exploration and asked why some of that computing power was not used to solve downstream and refining problems as well. His attitude was, "Let's try it and see how it works." The change at Shell, therefore, was driven by hardware, not software.

Researchers began looking for problems that modeling might be able to solve. Van Santen and the chemists who were willing began with extended Hückel calculations in 1986 and then rapidly moved to *ab initio* methods. Zeolites were chosen as the first catalytic system to study rather than metals because the structures were understood.

Then it became apparent that one could do homogeneous catalysis. Platinum-catalyzed polyethylene polymerization was a very important process, requiring a large ligand on the platinum atom. The same people working in the lab and product development worked very closely with theory. A spectroscopist who was a great inventor developed the new Ziegler-Natta catalysts at Shell. There was a critical mass of experience and knowledge around him that enabled him to jump into homogeneous catalysis with a running start. He asked van Santen to do calculations to test his hypotheses. He wanted to know the mechanisms to test his hypotheses, and eliminated hypotheses that were not scientifically valid. He was not looking for quantitative predictions, but wanted to enable conceptualization.

Van Santen left in 1988, but continued to work one day per week at Shell. From 1988-93, Shell had two people doing quantum mechanics, one doing homogeneous catalysis and one doing zeolites.

Berend Smit (now at the University of Amsterdam) started with Shell and then was sent to Berkeley, where he learned to do thermodynamics and diffusion calculations. Subsequently he returned to Shell and began to

do simulations. He had big success early on because he collaborated with information scientists who enabled large calculations on micelles.

Shell at that time had both heterogeneous and homogeneous catalysis modeling efforts going, but modeling never was completely accepted without dispute. There was no clear managerial support or clear value imparted. Not even within general research were simulations really predictive.

Smit left Shell in 1996. Shell stopped modeling work entirely in about 1997 because of a shift in research strategy to short-term development.

DISCUSSION

Research Strategies and Problem-Solving

In a number of companies like Shell and DuPont all of the current research is directly related to immediate business needs. There must be a clearly defined maximum of five years—no intermediate or longer-term research. Shell no longer does that kind of research, while it used to do so. For example, Prof. van Santen worked on the Fischer-Tropsch mechanism, but now Shell researchers work on how to scale manufacturing process and systems. The focus now is on engineering aspects, but they cannot commit themselves to the analysis of physical and analytical chemistry.

By extension, Prof. van Santen expressed the view that companies will rely much more on universities for such fundamental research in the future. He contends that "the cutback in modeling in industry is just turbulence, because industrial support for modeling in academia is increasing. Industry is convincing governments to set up theory centers in industry." In Europe, there is no shortage of funding for universities from industry. Nevertheless, people from industry come looking for new ideas. They do not want explanations of why things work. Rather, they want new things. Most research tries to explain what is happening and is not inventing.

Often, before managers will fund a modeling effort, they want to know what was done on the basis of predictions alone. That is the wrong question because usually success results from a team effort of experiment and modeling that solves a problem. Very often, several models could explain an experiment. Prof. van Santen noted, "we need to be able to use models to do a calculation accurately enough to make a prediction that can be measured, like climate is now. We cannot yet be certain enough to predict catalytic activity prior to synthesis. We cannot yet predict the world. That's why when a modeling group is separate, it gets murdered." The theorist should not take the risk of doing experiments but rather should collaborate with the best experimentalists.

In about 1994 Hoechst had a large lab in Frankfurt and a significant modeling group with much work in homogeneous catalysis. The company had five to six modelers, each of whom was part of a different project group. The modeling person was part of the invention project team and was included on patents. That group does not exist now, again because of the business climate. The business was split into several parts and smaller units at different sites. The corporate structure was chopped up.

In the lab at Haldor Topsøe (see CAMP Denmark report), there is no theoretical modeling group, but Nørskov spent 2-3 years there. Topsøe helped convince the government to fund CAMP as a major catalysis effort, and Topsøe benefits as do the Technical University and the students. Topsøe has a patent on a catalyst of some value, but the big payoff is the key insights on processes. It doesn't have experimental programs, so the company won't start modeling itself, but it creates great potential and opportunity for university programs.

DSM appears to be a different kind of company, in Prof. van Santen's opinion. DSM has the lab in the plant, and that means something. So for many years, there has been a big synergy between R&D and the plant. As a result, DSM did not have a dramatic change, like companies in which the central research lab is separate from plant and business units. Professor van Santen noted that when he visited DSM, it had a real freshness. The people at the bench had direct access to the head of the company and the research division. The head is

very active in the research content and activities. The head of research also spends one day a week there and invented an enzyme-catalyzed process. DSM researchers are innovative and imaginative, and they don't have a big problem with the "business vs. research" conflict. They don't see a conflict. Modeling is so entrenched that it is not vulnerable.

The History and Value of Commercial Software Houses for Catalysis

The history of commercial software for catalysis really began with Biosym and continued into its life as part of MSI (now Accelrys). The Biosym Catalysis Consortium began around 1988-89 with the principal focus being Richard Catlow's force field based methods to describe solid-state systems. This code was capable of doing beautiful predictions on structures slightly less complicated than zeolites. Under John Thomas's direction, Biosym funded students to develop code.

The first application was clear and well defined: X-ray diffraction and structural characterization. A major impact was made in the determination and elucidation of structures. Biosym's breakthrough in the early 1990s was that you could begin to do X-ray diffraction simulation in routine fashion.

Doing quantum chemistry on small silica systems, however, is still impossible. One needs to treat systems having 300 atoms. The commercial software has been very successful on analytical applications, but not on process applications. That is still a problem that has not been solved on the level needed.

Biosym's strength was the scientific expertise of its people and their interest in developing this expertise within the company. Admittedly, one technical shortcoming was that they were slow to develop density functional capability. Then at some point, they realized that the market would not sustain their catalysis effort. Their profit estimates from the mid-1980s did not hold up. Consequently, it is difficult to say how well their catalysis consortium will grow.

Professor van Santen, who also attended the GdR 12090 meeting at TotalFina, disagreed with a remark made there. In response to a panelist's comment that industry did not want to use academic codes, Prof. Juergen Hafner had interjected that all molecular modeling codes were really academic codes. It is true that many codes have academic origins, but usually the only way academic code is useful in industry is if people brought it with them and already know how to use it. Prof. van Santen commented that these programs must have a good interface and extension of inventions from academia, which is currently done in commercial software houses. Faculty may not think an interface is important because students run the code, but the interface is more than just window dressing. The Japanese have not had immense difficulty in improving the code for academics, but there are not enough resources to make a code useful for industry to use. This has also been a big problem for ADF. Plainly, there is an important role for software houses.

Both Prof. van Santen and Dr. Peeters believe that in some cases the government should fund the development of interfaces and manuals for academic code so that industry can use it. If it is development that is necessary but cannot be broadly supported by the market, then it should be government-supported. It is not appropriate for students or post-docs to develop such materials.

One way of addressing the need for good interfaces might have been the MSI SDK (Software Developers Kit), which allowed newly developed codes to be incorporated into MSI's interface. However, it came too late, and MSI was charging $25,000. People writing small programs did not want to pay that amount just for an interface. Now it's free, but it's too late.

European Development of Codes for Catalysis

The panel noted that there was a wave of belief in the mid-1980s that heterogeneous catalysis could be done, but that belief seemed to have died by the end of the 1980s. Many recent advances have originated in Europe, which leads the quantum chemical modeling of catalysis. From the panel's perspective, Prof. van Santen has provided much of the spark to initiate new work in heterogeneous catalysis today, along with Jens Nørskov with metals, Juergen Sauer and Richard Catlow in the chemistry community, and Matthias Scheffler in physics.

Prof. van Santen offered his own perspective on how these developments have progressed. The Planewave Program was originally developed by Mike Teeter at Corning to treat amorphous glasses (Corning bought Biosym in 1992 and MSI in 1995). He collaborated with Mike Payne at Cambridge, which led to CASTEP, now marketed by MSI. CASTEP was slow. CTEP was developed in a British consortium. Prof. van Santen's original focus was on clusters but then moved to periodic systems.

Prof. Juergen Hafner and postdoc Georg Kresse made a significant contribution by adding conjugate gradients for improved SCF convergence and soft pseudopotentials, both of which resulted in a dramatic reduction in cpu time required for these calculations. These developments were incorporated into the VASP program, and shown to be as accurate as gaussians for the total energies. VASP is free to academic groups in return for collaboration with Hafner. Van Santen was involved with VASP from the early Kresse days. FHI96MD was developed at the Fritz-Haber Institut in Matthias Scheffler's group. Bjork Hammer took the FHI code with him when he went to Denmark to Nørskov's group, which has developed its own codes as well.

Effective Academic-Industrial Interactions

Shell used to sponsor academic research, and the professor could then do what he wanted. Shell still does that to a small extent, but Prof. van Santen stated that he is really not interested in that. He believes in real, intense interaction. His industrial funding goes towards supporting specific students. TotalFina, as an example, selects the student. If industry makes the selection, the student is usually very good. In Europe, it is now possible to get a degree from two universities. The student gets experience in doing industrial triage in solving problems, and has to present his or her work to the companies.

Many of these relationships and funding have developed from consulting. Professors at Eindhoven need the university's approval to do consulting, as well as a decision on how much of the consulting fee they are allowed to keep. Professor van Santen gets paid a certain amount per day for consulting. If the company wishes to continue and expand the collaboration, an amount will be provided for a student and his or her research support, typically at $50,000 per year, although possibly for a period less than the entire PhD.

Prof. van Santen noted earlier that for many European companies all of the in-house research must be related directly to immediate business needs. An interesting aspect of industry's reduced level of in-house research is its increased dependence on university researchers. Top people leave, and companies do not replace them. As a consequence, they have expensive equipment that no one knows how to run. Prof. van Santen's group has been asked to run such equipment.

He pointed out that in Holland you have networks between business and universities, and that is crucial. Most important is capital investment by companies in computers and in academia. Most importantly, companies must invest in people. There are also needs for balance between methods and applications and for people to stay current in both methods and developments. Most of the groups he participated in had a specific objective to develop new methods. Now it may be more important to bring newer methods to the applications.

Distance is crucial for optimal interactions. If the distance is more than 150-200 miles, barriers arise. Prof. van Santen finds it easy for industry and academics (faculty and students) to meet in one day and travel back and forth. He thinks the United States should concentrate on several geographic centers (more than one), rather than establishing a single catalysis center. In Europe the focus is on physics and not on chemistry, with an emphasis on great accuracy for strong forces in small systems. This is a strength for methodologies but a weakness for simulations. That is why the simulation field is not that strong here, except for a few like Frenkel.

Looking to the Future: Needs and Targets

Professor van Santen hopes that he will be involved in applying computational chemistry to process design within five years. He wants to go through quantum to molecular mechanics to kinetic Monte Carlo. *Ab initio* methods will be used to calculate parameters to include the complexity of the environment with respect to temperature, pressure, and so on. Linking quantum mechanical calculations to predict kinetic macroscopic

phenomena is very important. With the VASP program, we are close to predicting rate constants for elementary reaction paths on catalysts. The problem is that the catalyst surface reconstructs. We need to be able to put in what happens over time with these complex phenomena, but it seems we are close. Also in the next five years, Prof. van Santen wants to begin doing the very-large-scale calculations of the computational materials science community but with catalysts.

Other developments and trends in Europe foreshadow future advances. Parinello is creating excitement doing simulation of acids in water, showcasing chemical problems. Also in Europe, electrocatalysis and fuel cells are a big growing area. For example, DaimlerChrysler is investing a big piece of its $23 billion research investment into fuel cells. Dr. Peeters pointed out the need to find replacements for different polymers like PVC, including whole new sets of additives for these polymers, analogous to plasticizers for PVC.

Dr. Peeters also remarked on the work of many companies in life science and molecular dynamics protein modeling. Hybrid methods are making a big impact and will continue to do so. However, although everyone is using hybrid methods, there doesn't seem to be enough validation. The big concern is the breaking of covalent bonds and how that will be treated accurately. Presently, quantum mechanics is only applied to one active site, not in more places where it might be valuable. In checking the Charmm-GAMESS interface, modelers found that simple things like rotation barriers didn't always come out right. There are real problems yet with these methods.

Combinatorial chemistry related to pharmaceuticals is already important, but its use for catalysis development will be increasingly important. Professor van Santen and his group already do some of the information science part of these rapid testing programs.

An important need in analytical measurements is to unravel the spectra of complex systems for which simplified test compounds or systems are not available. The key will be to crack the van der Waals problem. Prof. van Santen asserted that DFT is wrong in van der Waals systems. How can we treat zeolites even with Car-Parrinello if we can't do van der Waals? For polymers and soft materials it is crucial. Dr. Peeters speculated that semi-empirical methods will come back for situations where they are well suited.

The panel asked for comments about the usefulness of conventional tools. Dr. Peeters believes that with more computer power, the current methods will continue to help us more and more in the future over the next 15 years. That is not to say that major improvements aren't needed. In particular, most electronic codes don't parallelize well, so they don't do it. Neither does the van Santen group, although they use parallelized dynamic and kinetic Monte Carlo simulations over time and temperature regimes. Both agreed on the potential importance of Order(N) methods that are being developed.

Further along these lines, Prof. van Santen noted that conventional techniques available today will not carry us far into the future because of the time and length scale. Coarse graining needs to be developed to link the quantum and macroscopic world. This is a problem that has not yet been solved.

There is a particular need to find transition states and local minima on complex potential surfaces. This needs to be done much better, since current methods (including VASP) are still quite inadequate. What Topsøe is doing on surface transition states is very approximate, and the potential surface is only minimally and inadequately explored. It is still at the level of chemical intuition to set up approximate transition states by hand. Yet this is key to doing kinetics and dynamics on the surface. How can we link to longer time and length scales if we cannot search the potential surface and transition states thoroughly?

There is also a big need for improvements in statistical mechanics at different temperatures to obtain reaction rates. Low frequency modes are a problem for large statistical mechanics calculations; there are too many degrees of freedom. Also, proton transfer has quantum effects at 100°C. How then do we treat water accurately?

Site: **Rhône-Poulenc Industrialisation SA (now Rhodia)**
85, Avenue des Freres Perret - BP 62
69192 Saint-Fons
France

Date Visited: 16 September 1999

WTEC Attendee: P.R. Westmoreland (report author)

Hosts: Dr. Caroll Vergellati; Senior Staff Scientist and Molecular Modeling Manager, CRIT; T: (33) 4 72 89 69 25; F: (33) 4 72 89 67 22, caroll.vergelati@eu.rhodia.com
Dr. Simon Firkins, Research Scientist, simon.firkins@eu.rhodia.com
Dr. Dora Spyriouni, Research Scientist (Marie Curie Postdoctoral Fellow at the time of the visit), theodora.spyriouni@eu.rhodia.com

BACKGROUND

Rhône-Poulenc (http://www.rhone-poulenc.com/) was a company in the midst of sweeping merger-related reformation. At the time of the visit, businesses of the Rhône-Poulenc Group encompassed pharmaceuticals, agricultural chemicals, industrial chemicals, polymers, and specialty chemicals. The visit was made to its corporate research center.

In 1998, Rhône-Poulenc's chemicals component had been set apart as a wholly owned subsidiary, Rhodia (http://www.rhodia.com/). Rhodia has worldwide operations in fine organics (pharmaceutical ingredients, aromas, phenols, acetone), consumer specialties (food additives, detergents, personal care products), industrial specialty chemicals (silicones, silicas, latex), nylon fiber and plastics (second largest in the world), and services and specialties (membrane separations, cellulose acetate, catalysts, rare earths). In 1998 it employed 23,500 people and had sales of $5.5 billion.

During the second half of 1999, the life-sciences parts of Rhône-Poulenc and Hoechst (http://www.hoechst.com/) were merging into a new company, christened Aventis (http://www.rhone-poulenc.com/aventis/). On a 1997 basis, the components that formed Aventis employed 95,000 people and had sales of $20 billion, about 42% in Europe. Pharmaceuticals were 72% of 1997 sales, and agriculture products accounted for the balance.

At the same time, the two companies' chemicals businesses were being spun off completely as individual companies, Rhodia and Celanese AG (http://www.celanese.com/). The Hoechst components that will make up Celanese had 15,000 employees and $5.5 billion in sales in 1997. Its principal products are acetates, oxo and amine chemicals, acrylic acid, acrylates, oxo and amine products, catalysts, and polymers.

R&D ACTIVITIES

The meeting was held with the three members of the group that have conducted polymers modeling at Rhône-Poulenc. They moved into Rhodia in 2000, but the details and structure of the changes had not been announced at the time of the WTEC visit.

The group's activities had grown out of Dr. Vergellati's experimental background, complemented by molecular modeling experience during his post-doctoral research in Canada. He arrived in 1987 and began doing X-ray diffraction analyses of crystalline polymers. Using academic codes he had brought with him, he predicted certain of the polymer properties, which had a persuasive impact on his colleagues' interest in the area. In the early 1990s, a special group was formed, and Rhône-Poulenc became an original member of Biosym's Polymers consortium. Dr. Firkins arrived in 1991 with a theoretical chemistry background, having worked with Ian Hilliard at Manchester in developing the GAMESS-UK code. Coming from the group of Doros Theodorou, Dr. Spyriouni joined in 1998 as a post-doctoral associate under a two-year Marie Curie

Research and Training Grant. Her project was to be on molecular and meso-scale modeling of controlled release of bio-active molecules from polymers, emphasizing agricultural chemicals.

Table B1.9
Technical Summary of Polymer-Related Molecular Modeling at Rhône-Poulenc

Past and Present Applications	Important Emerging Applications	Theoretical Approaches and Software	Hardware
Interpreting X-ray diffraction data Predicting bulk properties of crystalline homopolymers Predicting mesoscale-based properties of polymer blends—morphology and evolution at temperature and in shear Polymers in solutions; gels Surface behaviors of coatings and adhesives Effects of inorganic fillers on polymer properties Homogeneous and heterogeneous catalysis NMR simulations	Properties for process design Stress-strain behavior of polymers Properties of semi-crystalline polymers	Molecular simulations of polymers mesoscale modeling MSI codes primarily; Gaussian codes in the past but not at present; Chem X originally	Workstations; no parallel computing used at present

Over time, a wide range of polymer-related modeling has been carried out. In French academia and industry, the scaling methods of de Gennes and others have had a strong influence on the study of polymer behaviors, yet molecular modeling is also proving quite useful. The emphasis has been on predicting and correlating polymer properties, but chemical reactions and surface or interface behaviors have also been studied. In the mid-1990s, there was a strategic move within the company toward specialty products like new additives.

There have been big successes, some of which hosts could not discuss. One success story was development of a new anti-scratch additive for polyurethane coatings. The project required high reactivity with a certain component of the paint but medium miscibility with solvent. The Flory interaction parameter χ was found for a number of molecular architectures, and relative reactivities were calculated with quantum chemistry. Without molecular modeling, the project would have been impossible.

In the biggest success (an unspecified application), a small number of calculations was first made to predict reactivity of a set of molecules. Internal experiments were carried out, and the predictions of reactivity were found to match. Based on this agreement, calculations were made on a large number of molecules. The best candidates were identified, synthesized, and tested.

It is important to note how valuable a negative result can be. In one case, a project was halted because calculations indicated the goal was not feasible. A Lycra-like material based on nylon was sought unsuccessfully. Six months of molecular modeling indicated that strong internal cohesion of the thermoplastic was necessary but could not be achieved with the nylon.

DISCUSSION

At present, the group relies exclusively on the MSI suite of programs. Group members were happier with the earlier, more collaborative days of the Biosym consortium, yet CAST and the COMPASS force field are

quite useful. Although they have used the Gaussian electronic-structure code in the past, they aren't using it at present. One issue is that they don't have any of the graphical user interfaces for Gaussian. Company policy is to use only commercial codes, so no academic codes are used.

Applications are developed internally, without academic collaboration. This has been partly because of the reluctance of academic purists to become entangled (seeking the best absolute number can be the enemy of accepting the good-enough number). Another reason is the strong influence of classical modeling of polymers among French academic circles. Rhodia is more inclined to collaborate through European programs, much as BASF and Unilever do. Also, beyond interactions with the Polymers Consortium, researchers attend some meetings (once every year or two), participate in CECAM workshops, and publish. Publication is more feasible if work is patented or if it didn't work as planned. If it didn't yield a sufficiently significant result, an academic case study is feasible. There have also been internal efforts to share methods and approaches, notably an internal global meeting two years ago on company computational methods that were in use.

CONCLUSIONS

Molecular modeling may not yet be perceived as a necessity for Rhône-Poulenc or its successor company, Rhodia. However, it has demonstrated its value and has convinced the Rhône-Poulenc R&D management of that value. There is increasing demand by the company researchers for results obtained by modeling, and the prospects for growth are good.

Site: **Royal Society of Chemistry Molecular Modeling Group**
Site: **Imperial College of Science, Technology and Medicine**
South Kensington Campus, Exhibition Road, London SW7 2AZ, UK

Date Visited: 13 September 1999

WTEC attendees: S.C. Glotzer and P. Vashishta (report co-authors), A.M. Chaka, P.T. Cummings, K.K. Irikura, P.R. Westmoreland

Participants:
Dr. Elizabeth Colbourn (coordinating host), Oxford Materials Ltd.; Chair, Royal Society of Chemistry (RSC) Molecular Modeling Group
Dr. Nicholas Quirke (coordinating host), Professor of Physical Chemistry, Head of Computational and Structural Chemistry, Dept of Chemistry, Imperial College
Dr. John Brophy, RSC
Dr. Darren Flower, Group Leader, Molecular Modeling and Bioinformatics, Edward Jenner Institute for Vaccine Research
Dr. George Jackson, Dept. of Chemical Engineering, Imperial College
Dr. John Kendrick, Wilton Research Center, ICI plc
Dr. Mario Moustras, RSC
Dr. Henry Rzepa, Dept. of Chemistry, Imperial College

Dr. Elizabeth Colbourn, Chair of Royal Society of Chemistry Molecular Modeling Group and Professor Nicholas Quirke, Head of Computational and Structural Chemistry, Department of Chemistry, Imperial College welcomed the WTEC group. The meeting was held at Imperial College at the South Kensington Campus.

There were two parts to the meeting. In the first part, the WTEC panelists introduced themselves and described the aims of their mission, while the UK participants described their backgrounds and research interests. The second part consisted of a broad discussion of past, present, and future of molecular and materials modeling in the UK.

BACKGROUND

Research Interests of Workshop Participants

Liz Colbourn was at ICI for about 15 years (leaving there about 7 years ago) where she established and led the Wilton modeling team. Her main areas of expertise lie in modeling catalysis and polymers using quantum chemistry, static lattice techniques, and molecular modeling. For a number of years, going back to her days in ICI, she has publicized the industrial benefits of modeling, both to industrialists and to academic groups. In 1993, she left ICI to start a small company, Oxford Materials, whose main product is PC software for polymer coatings, polymerization and particle packing. Oxford Materials (http://www.oxmat.co.uk) commercializes academic simulation tools and sells the resulting packages to larger companies. She is also the Chair of the Molecular Modeling Group of the Royal Society of Chemistry (RSC), which is part of the RSC's Industrial Affairs Division. In 1999, she co-founded Intelligensys Ltd. (http://www.intelligensys.co.uk/), a company focused on the use of artificial intelligence in product formulation.

John Kendrick was formerly at ICI (~20 years), and is now at Huntsman. At ICI he developed in-house expertise and then applied it to problems to add value. These include heterogeneous catalysis, polymer and material properties from quantum chemistry to finite element methods. ICI went through reorganization. There is no longer a central modeling function at ICI. The remaining modelers are now integrated into applications areas. Modeling methods are being applied vigorously (QSPR, polymer properties, catalysts), but long-term research is being neglected at ICI.

Darren Flower was educated at Imperial College (BSc) and the University of Leeds (Wellcome Prize PhD). He joined Astra (formerly Fisons Pharmaceuticals) in Loughborough as part of its molecular modeling group supporting projects in chemistry and biology. His work concentrated on most areas of molecular modeling, bioinformatics, and chemoinformatics. In 1999, he left Astra to set up a group in bioinformatics and molecular modeling at the Edward Jenner Institute for Vaccine Research (http://www.jenner.ac.uk), which is co-funded by the BBSRC, Medical Research Council, Department of Health, and a donation from GlaxoSmithKline. His work combines collaborative support for immunologists working at the Edward Jenner Institute with investigations into fundamental problems in immunology and vaccine design from a chemical theoretical perspective. He is also the secretary for the Royal Society of Chemistry's molecular modeling group.

John Brophy was formerly Corporate Research Coordinator at British Petroleum (BP) and General Manager at BP Chemicals. They used molecular modeling to assist in developing catalysts and new products, e.g., brake linings, polymerization catalysts, corrosion and scale inhibitors, which were all successful applications. Now he consults independently and also works with the RSC on industrial affairs and science policy, providing input to government departments on R&D policy.

The UK government's Foresight Programme was created to guide government investment in science and engineering R&D. Details can be found at http://www.foresight.gov.uk/. This program, which began in 1994, is the first one in which the UK government has sat down with academics and industry to try to structure and focus fundamental research (similar to the American "Vision 2020"). There are about 15 industry sector reports available from the first cycle of Foresight, and a further 12 were published in December 2000. All of these are available on the Web site.

Mario Moustras works in the scientific affairs department of RSC with John Brophy on Foresight activities. The RSC has 46,000 members worldwide, divided into six scientific divisions. He looks after the Industrial Affairs Division (14,500 members), which is comprised of 16 sections, including the Molecular Modeling Group. In addition, he also looks after other Foresight activities, such as research networking activities, scientific future for chemistry, and reports for materials chemistry.

Nick Quirke is professor of physical chemistry in the Department of Chemistry at Imperial College and head of the computational and structural chemistry research section. Prior to coming to Imperial College, he spent time at British Petroleum and Biosym. He joined BP in 1984 where he introduced statistical mechanics into BP's research effort. After 10 years he had about 10 people with funding from corporate and from business divisions and was able to make significant contributions to BP. Subsequently, BP went from being a very diversified company to being very focused, considerably reducing the value of long-term research within BP. After leaving BP, he worked in the Paris office of Biosym for a couple of years; this was a contract research model for the software company to sell research to customer companies. For the past four years he's been a full-time academic; his personal group consists of five postdocs and four graduate students. His group is supported by the research council EPSRC (60%) with about 40% industrial funding (largely from French and American companies). The group's work is a mix of long-term, fundamental research and applications to important problems. Current areas of research in his group are colloid chemistry—applications of statistical mechanics and simulations to "how colloids interact and phases they form"; nanoparticulates, where particles are not much larger than the solvent molecules they are in; and wetting properties. Other areas of research include applications to self-assembly and nanotechnology and adsorption and transport in porous media and mixtures in confined geometries on nanoscale. Transport in very small pores is still an unsolved problem. Prof. Quirke's group also is modeling electron trapping and transport in polyethylene (PE). PE was invented at ICI as a high-tension material for electrical insulation. PE-based materials break down over time, but it is not clear why.

Henry Rzepa is a faculty member in the Department of Chemistry at Imperial College. His background is in synthetic chemistry. Rzepa's work involves wavefunction-based molecular modeling (including density functional theory) for the design of synthetic reagents ("organic" is now passé in chemistry in the UK). He is inventing and engineering new reactions—using Gaussian etc. Wavefunction-based molecular modeling is being applied to cavity-based synthetic templates. Also, he and his colleagues are working on self-assembling molecules, cavity-induced chemistry, and development of Web-based strategies for disseminating information—XML and applications to cheminformatics and chemical resource discovery.

George Jackson is a faculty member in the Department of Chemical Engineering at Imperial College. He has a chemistry PhD (in statistical mechanics, not molecular modeling) from Oxford University (Rowlinson) and did his postdoctoral research with Keith Gubbins at Cornell. Among chemical engineering departments in the UK, this one (Imperial College) is closest to the American type: about one-half the faculty are traditional ChEs and the other half are chemists, physicists, and mathematicians. It's unique; all other chemical engineering departments in UK are very traditional. His research is in the areas of equations of state, complex fluids, phase equilibria of surfactant, polymeric, and liquid crystal systems. In the chemical industry, most of the cost is associated with separations, the theory of which is neglected in the UK. He wants to go beyond the heavily parameterized empirical models such as UNIFAC and UNIQUAC. He has been using statistical mechanics for the past 10 years to produce equations of state (EOS) that can be used by engineers. Techniques used involve perturbation and density functional theories. He has developed an equation of state (augmented van der Waals) to incorporate non-typical features like non-sphericity or strong associations or network formation, reactions, etc. It has been applied to refrigerants and interfacial energies, and phase diagrams can also be computed. He co-developed the SAFT method. ICI, BP, and Unilever provide funding, which drives the direction of the work: strongly associated fluids such as water, HF, formaldehyde, long-chain hydrocarbons, BF_3-alcohol mixtures, and refrigerants. For many properties, traditional statistical mechanics methods do as well as expensive molecular dynamics simulations. Some other successful applications include distillation of high-purity acetic acid (<200 parts per billion impurities) and phase behavior in oil extraction fluids at high temperature and pressure. SAFT method is his most cited paper. New grants are available in the UK to bridge communities. Most funds are going to catalysis research presently.

Other Activities

Some RSC committee members who were not at this WTEC workshop at Imperial College included Scott Owens (British Nuclear Fuels or BNFL), Kate Burt and Richard Lewis in pharmaceutical (Pfizer and Lilly), and Peter Coveney (formerly at Schlumberger, now at Queen's Mary College).

Another molecular modeling group in the UK is the Molecular Modeling Working Party (MMWP). It has 60 members with scientific background in computational chemistry. The MMWP meets at King's College and is directed by M. Robb. MMWP is mostly academic, in contrast to RSC's group, which is more industrial.

DISCUSSION

What is the perceived importance of modeling in industry? There are enlightened managers, but it is hard to answer this question. As new young researchers enter industry, they are more comfortable with modeling, and so companies become more comfortable.

Validation is a continuing issue. There are two kinds: validation of hard numbers and validation of trends. Both are basically left up to the individual researchers. The companies developing and/or using software are doing it on their own—there is no organized validation.

Equation of state (EOS) modeling and theory are very important—a critical need in industry. By far the most important thermodynamics in modeling is based on statistical mechanics. Process design packages like ASPEN still have very primitive inputs to which complex behavior cannot conform. These packages need to be more sophisticated. Even theory from the 1970s would be an improvement. Add "separations" to the list; it goes beyond simple vapor-liquid equilibria. In chemical engineering, global optimization is great but molecular sophistication is not. Industry needs a technique fast enough to run on a personal computer. The community needs better packages for separation processes, e.g., where reaction and distillation occur simultaneously, or when crystallization is involved. In chemical engineering, the constitutive relation is the equation of state. We need to generate a mathematical tool that people can use easily—reduction of parameters. But is an equation of state derived from statistical mechanics really needed, or is lots of data plus extrapolation sufficient? This is what's done often in process simulation—they rely on databases. An EOS is needed when real data are not available, e.g., gas-adsorption methods; information from simulation can be used to fill the gap in databases, which is then used for prediction. Other needs, in conjunction with

experiment and simulation, include (1) refinement of crystal and NMR structural data for biomolecules, (2) deconvoluting adsorption data of gases in porous materials.

In several pharmaceutical companies there are substantial-sized modeling groups (10-20 people) working on drug design. The question has been asked: "How many drugs have been designed using modeling?" There are some examples, like at Vertex and Merck. However, no company will produce a drug that has been solely virtually designed—the process still has to go through laboratory testing. A reasonable point of view—one put forth in the UK by Graham Richards—is to recognize that it is no more sensible to ask, "what drug has been developed by NMR?" Modeling is currently considered an integral part of pharmaceutical development, and it may contribute at all stages, but not concretely or identifiably. This is true not only in pharmaceuticals but also in materials.

Molecular modeling is gaining credibility within the pharmaceutical industry, but molecular modelers seem to be virtually absent from UK patent lists. Modeling approaches are now more successful in a business sense since they give results you can use in the laboratory. Old-style modelers are now moving toward bioinformatics and cheminformatics. It is understood in the pharmaceutical industry that modeling is a key part of the whole R&D process.

Convincing experimentalists that modeling has something to contribute is difficult. It is important to convince them you're not a threat, that you'll enhance their work, not eliminate it. Your role is to help simplify the complexity. You can computationally realize the experimentalist's conceptual model.

Modelers should be part of the team from the very beginning. Modeling is merely a simplification of a complex process. Just asking the simple questions is useful. Modeling provides a common language within which to discuss the problem. Modeling provides the "aha" effect (e.g., "The whole chain is moving!! The NMR guys only told us the rings were flipping!") Models allow one to disseminate information. "You don't get out of bed without a model."

At the beginning, getting industry to buy into molecular modeling was a sales job. Then once results were there, it was easier. Still, it really hinged on individuals within the company, both on the seller's side and the buyer's side. Pictures aren't sufficient for satisfying managers anymore, although they did at the beginning. Prediction—in the end, that's how we're judged as modelers. It's only the tip of the iceberg, yet it is all that management sees. On the other hand, if modelers sell their stuff based solely on prediction, then they'll be in trouble.

Companies don't like to describe either achievements or failures in detail. Molecular modeling has reached the same level of (in)security as other research areas. Prediction is the criterion for judging molecular modeling efforts, but straight prediction is really not very important.

In the United States, the semiconductor industry is way ahead in recognizing the importance of fundamental research. Outsourcing is the latest strategy being implemented by industry to accomplish long-term research, but this will backfire, since it's hard to work out intellectual property, and also because there's no one left in the company to know how to outsource.

Kendrick and Colbourn reflected on their ICI experiences. Modeling can provide concepts and language that provide a new perspective on a problem. "It's difficult to separate what's happening in research from what's happening in modeling." The new, younger people are more comfortable with modeling. In 1991 at ICI a cost-benefit analysis was done for the molecular modeling group, with very favorable results. Some answers included that 6 months of modeling saved 6 years of project time. Still, how do you put a dollar value on the impact of modeling?

Quirke reflected on his MSI experience. Simulation software companies like MSI have played a big role. Their consortia are important. It provides publicity for molecular modeling in general. The Compass force field in MSI is excellent. That's one thing they did well. Outsourcing to universities by industry will continue to grow. But small companies will be associated with them (the university people). The easy stuff with be routine and will be done in-house in industry.

Where will commercial application of modeling be in 15-20 years in the UK? The answer depends on the state of research in the UK at the time. Colbourn remarked, "It will be a routine and necessary part of business. But I don't know if that business will be in the UK!" Certainly there will be more outsourcing to academics and small companies associated with universities. But software companies won't be doing the research themselves. Modeling will be embedded in the undergraduate curriculum, and software will be very cheap and easily accessible on PCs. It inevitably will become a normal, necessary part of business in the long term, with the really complicated problems being outsourced, maybe.

MOLECULAR MODELLING GROUP—OFFICE HOLDERS AND COMMITTEE MEMBERS (1999 AND 2000)

Dr. Elizabeth Colbourn – Chair
Intelligensys Ltd., Belasis Business Centre, Belasis Hall Technology Park, Billingham, Teesside TS23 4EA; tel: + 44 (0) 1642 343411, fax: + 44 (0) 1642 714305, Email colbourn@intelligensys.co.uk

Dr. John Kendrick – Treasurer
13 Castle Close, Middleton St George, Darlington, DL2 1DE; Email John_Kendrick@huntsman.com

Dr. Darren R Flower – Secretary
The Edward Jenner Institute for Vaccine Research, Compton, Berkshire RG20 7NN, tel +44 (0) 1635 577954, fax +44 (0) 1635 577901, Email darren.flower@jenner.ac.uk

Dr. John Brophy, Royal Society of Chemistry, Burlington House, Piccadilly, London W1V 0BN, tel +44 (0) 20 7440 3307, fax +44 (0) 20 7437 8883 Email Brophyj@rsc.org

Dr. Robert Docherty, Head of Materials Sciences, Pharmaceutical R&D, Pfizer Global Research and Development, Sandwich, Kent, C&13 9NJ, tel +44 (0) 1304 642068, fax +44 (0) 1304 653909, Email Robert_Docherty@sandwich.pfizer.com

Dr. Paulette Greenridge, Thrombosis Research Institute, Emanuelle Kaye Building, Manresa Road, London SW3 6LR, tel +44 (0) 20 7351 8330 ext 3622, fax +44 (0) 20 7351 8324 Email paulette@tri-london.ac.uk

Dr. Richard A Lewis, Computational Chemistry, Eli Lilly and Co Ltd., Lilly Research Centre, Erl Wood Manor, Windlesham, Surrey GU20 6PH, tel +44 (0) 1276 653236 Email lewis_richard@lilly.com

Dr. Steve Maginn, Cambridge Crystallographic Data Centre (CCDC), 12 Union Road, Cambridge CB2 1EZ, tel +44 (0) 1223 762534, fax +44 (0) 1223 336033 Email maginn@ccdc.cam.ac.uk

Dr. Scott Owens, British Nuclear Fuels Ltd., R202 Rutherford House, Risley, Warrington, Cheshire, WA3 6AS, tel +44 (0) 1925 832178, fax +44 (0) 1925 833561 Email slo1@bnfl.com

Professor S. Ramdas, Schroedinger Inc., 3 Lakeland Drive, Frimley, Surrey, GU16 5LD, tel +44 (0) 1276 684507, fax +44 (0) 1276 507181 Email ram@ramdas.fsnet.co.uk

Dr. Antony J. Richards, Molecular Simulations Inc., 240-250 The Quorum, Barnwell Road, Cambridge CB5 8RE, tel +44 (0) 1223 413300, fax +44 (0) 1223 413301 Email arichards@msicam.co.uk

Andrew Sparkes, Tripos UK Ltd., 13 Shenley Pavilions, Chalkdell Drive, Shenley Wood, Milton Keynes, Bucks, MK5 6LB, tel +44 (0) 1908 507333, fax +44 (0) 1908 507444 Email asparkes@tripos.com

Dr. Nick Tomkinson, Physical & Metabolic Sciences, AstraZeneca Charnwood, Bakewell Road, Loughborough, Leics, LE11 5RH, tel +44 (0) 1509 645824, fax +44 (0) 1509 645576 Email nick.tomkinson@astrazeneca.com

Prof. Paul Barnes, Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, tel +44 (0) 171 631 6817, fax +44 (0) 171 631 6803, barnes@gordon.cryst.bbk.ac.uk

Dr. Catherine Burt, Computational Chemistry, PC 557, Pfizer Central Research, Sandwich, Kent CT13 9NJ, tel +44 (0) 1304 648464, fax +44 (0) 1304 658463, Email burt@pfizer.com

Dr. Peter V Coveney, Room Chem9l6, Department of Chemistry, Queen Mary and Westfield College, Mile End Road, London El 4NS, tel +44 (0) 171 975 5026, fax +44 (0) 181 981 8745. Email p.v.covney@qmw.ac.uk , URL http://www.chem.qmw.ac.uk/

Professor Nick Quirke, Professor of Physical Chemistry, Head of Computational and Structural Chemistry, Department of Chemistry, Imperial College, South Kensington, London SW7 2AY, tel +44 (0) 171 594 5844, fax +44 (0) 171 594 5850, Email n.quirke@ic.ac.uk

Vanessa Spedding, Scientific Computing World, IOP Publishing Ltd., Dirac House, Temple Back, Bristol BS1 6BE, tel +44 (0) 117 929 7481, fax +44 (0) 117 930 1178. Email scicomp@ioppublishing.co.uk

Site:	**SmithKline Beecham Pharmaceuticals (now GlaxoSmithKline)** **New Frontiers Science Park (North)** **3rd Ave.** **Harlow, CM19 5AW** **United Kingdom**
Date Visited:	13 September 1999
WTEC Attendees:	P. Kollman (report author), R. Miranda, K. Morokuma, E. Stechel, A. Tsao
Hosts:	Colin Edge (Colin_M_Edge@sbphrd.com) Martin Saunders

BACKGROUND

SmithKline Beecham (http://www.sb.com/) is a large, multinational pharmaceutical company, one of the world's largest, with significant research facilities in a variety of locations and the largest locations in the United States and United Kingdom. Subsequent to the site visit, it merged with Glaxo Wellcome to become GlaxoSmithKline.

R&D ACTIVITIES

The computational modeling group at SmithKline Beecham (SKB) consists of nine people, of whom Colin Edge is the group leader and Martin Saunders is a member. Computational modeling is part of the Computational and Structural Sciences Section, which includes X-ray, NMR, mass spectrometry, synthetic methods, etc. There are five other sections at the site—Med. Chemistry I to IV and Combinatorial Chemistry—in the general area of discovery chemistry. Edge has a background in quantum chemistry, and Saunders is trained in semi-empirical QM, molecular mechanics and solvation effects. The headcount of the group (nine) is as low as it has been, and there has been a headcount freeze recently. Still, the number of modelers compared to the number of PhD researchers on the site (250) was thought to be reasonable.

The commitment of SKB worldwide to genomics has changed the way targets are chosen for design within the company. The organization is a matrix one, so the computational chemists are part of the project teams (horizontally) as well as the computational chemistry group (vertically). There are certain areas of expertise within the group—Frank Blaney is an expert in modeling various G-protein coupled receptors, while Saunders focuses on enzymes, and kinases—but most have to be "jacks of all trades" when part of a project team. Because of the large number of targets developed through genomics, individuals are involved in as many as seven targets at once. Each project is reviewed twice a year. It was felt that there was good communication among the modeling group and between the group and the chemists, so people knew where and whom to seek for various expertise.

The group has a 32-processor O2000 from SGI, approximately 40 SGI Indigos, 8 processor SGI Challenges, and a truckload of PCs. They use a significant amount of commercial code and also buy codes from universities. They are not hardware-limited in what they do. The budget for software is about $1 million/year and includes licensing SYBYL, QUANTA, Insight II, AMPAC, Gaussian, Spartan, Daylight, Unity, Catalyst, and ISIS.

The number of targets with high-resolution structures has increased from about 25% to 50% in recent years. In addition, Blaney has done extensive modeling on many G-protein coupled receptor systems, where one knows the structure to relatively low resolution. Specifically in the last two years, about 45% of the targets had structures or well-defined homology models and the group worked on 26% of the 300 targets proposed. About 32% of the targets worked on are G-protein coupled receptor targets.

There are three protein crystallographers on-site and eight in the United States, as well as three macromolecular NMR researchers on-site. There is a group in chemical informatics and bioinformatics that

has a headcount of 40 and does a significant amount of analysis of physical property correlation. There is a cheminformatics group of 4-6 headed by George Seibel. As noted above, there is a combinatorial chemistry group on-site, and the computational chemistry group helps them in library design, mainly using scripts within the Daylight tool kit. They put filters in to make sure the ligands meet the Lipinsky criteria for good physical properties.

DISCUSSION

The discussion noted the historical trends in expectations for computational chemistry. In 1987, there were unrealistic expectations; in 1990, a trough of despondency; and now there are reasonable expectations of what it can contribute. There are still some non-believers but not as many; the expectations are reasonable, and computational chemistry is recognized as part of the drug design process. Because of the use of genomics, the computational chemist often can begin a project before the chemists are involved because the structure of the protein can be available before synthetic efforts are undertaken.

The computational chemistry group also does education internally for chemists interested in using the modeling tools themselves, initiating a forum called "Modeling Chemists," giving training in a subset of the computational tools such as SYBYL and QUANTA (maybe TSAR and CATALYST). The number of chemists who have been trained is in the range of 30-40. This approach has been used in the United Kingdom only, not at the U.S. site.

Publishing is apparently valued for promotion, but not for day-to-day work. Computation reportedly made a significant contribution to the next blockbuster SKB drug, a broncodilator PDE4 inhibitor. The computational chemists are considered for listing on patents, just like other members of the project team.

The group members have very useful outside interactions. They have some in-house SANDWICH students, one per year in computational chemistry, and support one or two CASE students per year. There is no support of postdocs outside, but there are 10 postdocs total in all the research groups in discovery chemistry on site. They would be amenable to contracting out some aspects of modeling if the price is right.

The group collaborated previously with IBM Winchester, in what became ANACONDA, to represent molecular fields. SKB has agreements with Cambridge Xstal database and ISIS to have access and field test databases and software. They plan to join the Klebe consortium to build a better scoring function for molecular interactions. They plan to work with EMERCK to develop a database of ligands in protein binding. They generally have not found consortia particularly useful with the exception of the one organized by Klebe, where they clearly will get information not available just from their own data.

The main challenge for the future noted by the group was the need for better scoring functions for macromolecule-ligand interactions, including protein flexibility in the calculation.

Site:	**TotalFina (now TotalFinaElf)** **Establissement de Normandie** **European Technical and Research Center** **B.P. 27—76700 Harfleur** **France** **Tel: (33) 35 55 11 11, Fax: (33) 35 55 14 80** **http://www.totalfinaelf.com/ho/en/**
Meeting site:	Joint visit with GdR at TotalFina Raffinage, Le Havre, France
Date Visited:	15 September 1999 (Joint visit with IFP and GdR 12090 project)
WTEC Attendees:	A.M. Chaka (report author), P.T. Cummings, K.K. Irikura, P. Vashishta, P.R. Westmoreland
Host:	Dr. François Hutschka, francois.hutchska@total.com

BACKGROUND

TotalFinaElf (http://www.totalfinaelf.com/ho/en/) is the largest oil company in Europe. TotalFina, the company's name at the time of the site visit, was a product of the 1998 merger of Total of France with Petrofina of Belgium. A merger with Elf Aquitaine, also of France, was announced in September 1999 at the same time as the site visit, so the site report does not reflect any consequent changes.

TotalFina has a central corporate R&D organization, the European Technical and Research Center (CERT—Centre Europeen de Recherche & Technique) located in Harfleur, France, employing nearly 300 researchers, engineers and technicians. There is a frequent exchange of staff from central R&D to the commercial and manufacturing groups to encourage communication and interaction.

RESEARCH ACTIVITIES

The overriding mission of CERT is to provide the refining segment of the company with the expertise needed to get the best value out of crude oil. This consists principally of improving existing refinery products, developing new ones, and designing production infrastructure. Examples of downstream products are formulated lubricants, fuel oils, solvents, bitumen, paraffins, and micro-waxes.

Current R&D is largely driven by refining issues arising from the need to obtain a consistent quality of products from increasingly heavy available crude oils whose quality and characteristics vary considerably with the source, and from increasing environmental concerns.

Catalysts are key to most refining processes and are a major focus of R&D. Most of the modeling work is focused on understanding the mechanisms of homogeneous and heterogeneous catalysis, and in calculating thermochemistry using *ab initio* methods.

TotalFina has an "open-door policy whose purpose is to keep research at the cutting edge of technology." This consists of forming research and development partnerships with international customers and suppliers, as well as with France's government-sponsored research foundation (CNRS) and the Institut Français du Pétrole (IFP). TotalFina frequently licenses technology from IFP. The company also funds graduate students and hosts interns from France's universities. CERT staff is also encouraged to teach at the country's various engineering schools.

DISCUSSION

The use of computational chemistry at TotalFina began in 1994 with the participation in the Biosym, now MSI, consortia (Catalysis and Sorption, Polymer, and Electronic, Optical, and Magnetic Properties of Materials (EOM)). The molecular modeling expertise of TotalFina is located at the CERT. Dr. Sylvain Clemendot began the modeling effort at TotalFina, but in 1998 he moved to the pilot plants where he is in charge of trouble-shooting problems.

Staffing for computational chemistry consists of one full-time researcher working primarily on refining-related projects. Additional work can arise from problems and/or questions in the exploration-production area or specialty chemicals (inks, coatings, adhesives, and paint).

Examples of applications are:

- Catalytic reactions
  - Hydrodesulfurization of refractory molecules on transition metal sulfides
  - Alkane transformation (cracking, isomerization, etc.) on zeolites
- Properties or reactivity of "isolated" molecules
  - Acidity, solvation energies, thermochemistry, etc.
- Characterization of molecules and/or periodic systems
  - Acidity of zeolites, IR or NMR spectra, etc.
- QSAR studies
  - Efficiency of corrosion inhibitors
  - Molecular properties like boiling points, etc.

TotalFina actively engages in external collaborations in molecular modeling. The company works closely with MSI through consortia, participates in a European research group (with CNRS, IFP, TUW) dedicated to the application of theoretical chemistry to catalysis (GdR 12090), and supports PhD students at many universities.

Software and Hardware Tools

- Quantum (SMOL3, CASTEP, VASP, GAUSSIAN 98)
- Classical (DISCOVER)
- Statistical (QSAR)
- Human: Neurons and knowledge (The most important!)
- Several Unix workstations (SGI)
- One parallel server (SGI O2000 with 16 processors) dedicated to time-consuming calculations

CONCLUSIONS

Molecular modeling has enjoyed many successes at TotalFina and is well funded with respect to hardware, software, and external collaborations. Francois Hutschka, however, has many more requests for his services than time available and finds it amazing that, given the company's resources, they have only one full-time staff member doing modeling.

Site: **Unilever**
Quarry Road East, Bebington, Wirral L63 3JW
United Kingdom
http://www.unilever.com/

Authors: P.T. Cummings and P.R. Westmoreland

BACKGROUND

The following report is based solely on publicly available sources.

Unilever is a transnational corporation specializing in home and personal care products and food products. It has manufacturing capabilities in more than 90 countries and sells its products in these and another 70 countries under hundreds of familiar names (such as Promise margarine, Lipton tea, the Calvin Klein cosmetics, Aim toothpaste and Dove soap). Unilever's specialty chemicals business (the National Starch and Chemicals unit) was sold to ICI in 1997 for $8 billion as a key component of ICI's revised corporate objectives. Sales of Unilever are around $45 billion/year with before-tax profits of around $5 billion/year.

Unilever spends $1 billion/year on research. It has six large research centers: Colworth (food); Edgewater, NJ (home and personal care); Vlaardingen, The Netherlands (food and laundry); Bombay; Shanghai; and Port Sunlight (laundry, hair deodorants, and household care).

MOLECULAR MODELING R&D ACTIVITIES

By external reputation, Unilever has one of the most active and successful MMS efforts in the world. Competitors report that Unilever employees have said that roughly thirty people are working in the area.

The Physical and Measurement Sciences Group (PMSG) located at Port Sunlight is under the leadership of Dominic Tildesley, who joined Unilever in this capacity in mid-1998 after a successful academic career in chemistry at the University of Southampton and London's Imperial College. A pioneer in molecular simulation research, he is also well known as co-author of the most widely used and cited monograph on molecular simulation (Allen and Tildesley 1987).

An unusual feature of Unilever is the substantive development of theory and codes, most prominently seen in the dissipative particle dynamics (DPD) method of Warren (Warrant 1998; Groot and Warren 1997; Groot and Madden 1998). DPD is a coarse-grained molecular-simulation methodology that can be used to study phenomena on much longer time scales than permitted by molecular dynamics. Unilever has published many of the seminal papers in the field, and a commercial DPD code is presently marketed by MSI.

Another activity has been molecular design of a peptide bacteriocide (Patel et al. 1998).

REFERENCES

Allen, M.P., and D.J. Tildesley. 1987. *Computer Simulation of Liquids*, Oxford, Oxford University.

Warren, P.B. 1998. Dissipative particle dynamics. *Current Opinion in Colloid & Interface Science* 3, 620-624.

Groot, R.D., and P.B. Warren. 1997. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* 107, 4423-4435 (1997).

Groot, R.D., and T.J. Madden. 1998. *J. Chem. Phys.* 108, 8713; http://www.msi.com/materials/cases/copolymer.html.

Patel, S., I.P. Scott, M. Bhakoo, and P. Elliot. 1998. Patenting computer-designed peptides. *J. Computer-Aided Molecular Design* 12, 543-556.

APPENDIX B.2. SITE REPORTS—JAPAN

Site: **Asahi Chemical Industry Co., Ltd.**
R&D Division
2-1, Samejima, Fuji-City
Shizuoka 416-8501, Japan

Date Visited: 10 November 1999

WTEC Attendees: R. Miranda (report author), P.R. Westmoreland, P.T. Cummings, S.C. Glotzer, K. Morokuma, E.B. Stechel

Hosts: Tomoyoshi Yamada, R&D General Manager, tel. 545-62-3047, fax 545-62-3039
Yasuhiro Ueshima, General Manager, Comp. Science Department, ueshima@cs.fuji.asahi-kasei.co.jp
Terumasa Yamasaki, Comp. Science Department, yamasaki@cs.fuji.asahi-kasei.co.jp
(Takeshi Aoyagi, Comp. Science Department, aoyagi@zoom.cse.nagoya-u.ac.jp is currently with the Doi Project)

BACKGROUND

Asahi Chemical (http://www.asahi-kasei.co.jp/asahi/), with sales of ¥1,172 billion ($11.0 billion) in the year ending March 31, 1999, is the largest chemical company in Japan. Asahi Chemical has over 29,000 employees (1999). Their core businesses are chemicals and plastics, housing and construction materials, fibers and textiles, and pharmaceutical, medical and specialty products. The basic chemicals such as methyl methacrylate, adipic acid, and styrene; fertilizers; and polymers (PE, PP, SA, ABS, PMMA, mPPE, etc.) were responsible for 32% of profits in 1999. Housing and construction products also yielded 32% of overall profits. Specialty products yielded 24% of profits— with pharmaceutical and medical products, chemicals for electronics fabrication, membranes, foods and liquors, and engineering services. Synthetic fibers, cellulosic and non-woven fabrics produced 12% of profits.

R&D ACTIVITIES

The Analytical and Computational Sciences Division is part of the Corporate R&D Section, which is directly linked to the company president. Drug design activity is carried out separately under the pharmaceutical and medical products division.

This division has 14 employees, five of whom perform molecular and materials modeling. Three of them are computational chemists and one is a solid-state physicist. Their expertise lies primarily in the application of *ab initio* electronic structure calculations (including DFT), force field-based molecular dynamics, and solid-state material modeling. Dr. Yamada has been with the company since 1973 and with the division since 1996. Dr. Ueshima has also been with this division for about four years. Mr. Yamasaki joined the company in 1985, and since then he has been with this department. He had two years of experience with Goddard (Caltech) in transition metal catalysis (GVB and force field). The computational sciences department had a maximum of 24 persons in 1996. The number of people in the department has since declined, but the primary reason is transfer of personnel to the information technology department and the pharmaceutical divisions for computational research work.

The goals of the computational science division are (1) to reduce R&D costs and development time, (2) to guide experimental activities with reliable predictions, and (3) to develop qualitative reasoning models such as QSAR (QSPR). Applications are directly related to the company's core businesses, primarily to medical products, polymers, new materials for electronics, and catalysts. Support for the research units involves about three projects per modeler, as well as computational training and consulting. In 1997 they undertook 36

projects, 12 of which were in computational chemistry. In 1998 they were involved in 41 projects, 14 involving computational chemistry.

Hardware and Software

The computational sciences department maintains and operates a Cray T94 (1 x 2 Gflops CPU, 512 MB memory), a Cray J932 (16 x 200 Mflops CPUs, 4 GB memory), and an Auspex centralized file server with 72 Gbytes and about 30 workstations (SGI, HP, Sun and IBM). For molecular and materials modeling, they use primarily Cerius2 (Materials, Inorganic, Polymer and Organic modules), Unichem, DGauss, Cognac (mesoscale simulation developed at the Doi Project), Gaussian, CASTEP and ESOCS (for band structures), Synthia (QSAR-based polymer property prediction), and MM/MD. For computational fluid dynamics simulations they use StarCD. They usually refer to standard structural databases (Cambridge and ICSD). For LSI process and device simulation, they employ mainly TSUPREM4 and MEDICI.

Project Selection and Examples of Successes

Computational modeling projects are originated at the request of corporate R&D divisions as well as business divisions. Projects are selected on the basis of level of importance for the company and the likelihood of success. The beneficiary divisions are charged a modest amount (10-15% of the actual cost) although that percentage may be increased in the near future, following company suggestions. The modeling group also takes up a small number of their own calculations in order to develop expertise for future projects.

One of the significant successes was the design of a single-site Zr metallocene olefin polymerization catalyst. Asahi's prediction preceded the experimental verification by about three years and led to a patent application. Results of the project included a ligand database, an associated QSAR and a prediction method. The algorithm produced lead catalyst formulations that were then optimized.

Their work has also led to significant advances in structural analysis of heterogeneous catalysts, combining force field, DFT calculations, and experimental information. In polymer design, they have developed a coarse-grained MD for polymer properties. They foresee increased use of mesoscopic simulation for polymer design and property calculation. They expect improved prediction of crystal structures by first-principles MD, as well as catalyst design by DFT and *ab initio* methods. The use of molecular modeling will also be increased in the next 5 to 10 years for electronics materials design, electric, optical and magnetic property prediction, and process simulation. Combinatorial methods and library design, QSAR and data mining will have larger applicability in catalyst and chemical development.

Immediate and Long-term Needs

Immediate needs were described as (a) mesoscale simulation for polymer properties, (b) *ab initio* electronic structure calculation for kinetics prediction in the presence of solvents or counterions, (c) interfacial property prediction, (d) library design and QSAR methodology, (e) molecular modeling software parallelized for personal computer clusters.

Long-term needs were seen to be (a) *ab initio* structure prediction for inorganic and metal oxide crystals of 100-200 atoms per unit cell, (b) prediction of polymer morphology and related properties, (c) prediction of impurity, point and mesoscale defects, (d) automatic construction and analysis of reaction mechanisms, (e) integration of theoretical and computational tools and databases, (f) software to perform all of the above in workstations, and (g) software to replace chemistry handbooks.

Theoretical chemists in the future will probably be developing useful algorithms for order-N (linear scaling) calculations of electronic structure, force fields, and statistical dynamics. Theory will lead to new concepts in chemical reactivity in terms of electronic structure. Electronic-based property descriptors and combinatorial approaches will become commonplace.

In the future, routine use of computational chemistry software by bench chemists will be common. Currently the reliability, quality, and price of software are unattractive to many non-specialist chemists.

Transfer of Science to Technology

Collaborations between academia and industry should occur at the planning and development of scientific programs. Asahi maintains active collaborations with Prof. Doi (Nagoya) and Goddard (Caltech). Industry has the role of helping to define attractive targets, and academia of providing innovation and next-generation scientific methodology. Government support at the initial stages is fundamental. Software houses should join the projects in a second stage, when need for code improvement and distribution becomes relevant. Asahi is a member of MSI's Polymers Consortium.

The transfer of science to technology is encouraged by defining goals that cannot be achieved readily in an experimental manner, for example, prediction of non-observable properties and species (transition states, reactivity indices, descriptors). Scientific outcome should be packaged (for example, as in the Gaussian program) to facilitate application.

CONCLUSIONS

Asahi Chemical has a mature computational sciences group that has acquired stability by means of regular interaction with experimental groups and effective interaction with company directors. They have established a good reputation by consistently contributing meaningful (value-added) results to long-term projects. However, they avoid research projects that lie beyond the capability of current theory and simulations. Prime examples of recent successes are (1) the design of polymers from structure-property relationships, (2) the discovery of novel homogeneous catalysts, (3) the structural analysis of heterogeneous catalysts.

The computational sciences group places major emphasis on the future availability of molecular modeling tools in the hands of the practicing chemist. They actively educate their lab-bench colleagues and expect in return a more meaningful and active cooperation.

In a diversified company such as Asahi, prioritizing has been key to the success of the computational group. In process chemistry, their current priorities lie in catalysis, thermochemistry, kinetics and structure prediction. In materials, they focus on strength, deformation, electrical, magnetic and electronic properties. In the future, they will concentrate on surfaces and interfaces. The Asahi computational group predicts it will also play a critical role in new discovery technologies, such as library design, diversity control, QSAR, and data mining.

REFERENCES

Yamasaki, T. 1995. Structure and lewis acid sites in alumoxane compounds. *Catal. Today* 23, 425.

Yamasaki, T., and W.A. Goddard III. 1998. Correlation analysis of chemical bonds. *J. Phys. Chem. A* 102, 2919.

Site: **"The Doi Project"**
Nagoya University
Department of Computational Science and Engineering
Graduate School of Engineering
Furo-cho, Chikusa, Nagoya, Japan 464-8603
http://www.stat.cse.nagoya-u.ac.jp/ (in Japanese)

Date Visited: 9 November 1999

WTEC Attendees: S.C. Glotzer (report author), P.T. Cummings, R. Miranda, K. Morokuma, E.B. Stechel, P.R. Westmoreland

Hosts: Dr. Masao Doi, Professor, Dept. of CSE, Nagoya University, 81-52-789-5633; Fax: 81-52-789-5831; doi@cse.nagoya-u.ac.jp
Dr. Toshihiro Kawakatsu, Associate Professor, Dept. of CSE, Nagoya University
Dr. Takashi Igarashi, Mitsui Chemical Inc.; Senior Research Manager, Doi Project
Takeshi Aoyagi, Asahi Chemical; Researcher, Doi Project
Hiroo Fukunaga, Fujifilm; Researcher, Doi Project

BACKGROUND

The project "Platform for Designing High Functional Materials," widely referred to as the "Doi Project," is a government-funded, university-led, four-year program to develop a computational framework for the simulation of multicomponent polymers from the nano- to the macro-scale. The project is being carried out by the Japan Chemical Innovation Institute (JCII) under the leadership of Professor Masao Doi at Nagoya University.

Implementation of the project was entrusted to JCII by the New Energy and Industrial Technology Development Organization (NEDO, http://www.nedo.go.jp/english/) under the Ministry of International Trade and Industry's (MITI) Program for the Scientific Technology Development for Industries that Create New Industries. NEDO, which was established in 1980 as a semi-governmental organization under MITI in fields related to technological development in Japan, works to coordinate the funds, personnel, and technological strengths of both the public and private sectors. NEDO's activities include such fields as development and promotion of new energy and energy conservation technologies, research and development of industrial technologies, restructuring of Japan's domestic coal mining industry, production of industrial alcohol, and restoring damaged coal mining areas. Of the roughly 30 projects within NEDO, the Doi Project is currently the only one dedicated to molecular and materials modeling.

Total funding for the Doi Project for 1998-2002 is $16 million. Of this total, JCII receives $4 million/year from MITI through NEDO and receives the remainder directly from the Ministry of Education. Most of the money is used to pay the salaries of the industrial participants (see below) and for lab construction and computer equipment. Salaries of the Nagoya University professors are not paid by the NEDO grant.

Project History

In 1989, a computational-chemistry-based industry working group entitled, "Innovative Chemistry by Computation" was formed within ASPRONC (Association for Progress of New Chemistry, a chemical industry association). The goal of the working group was to gather ideas and build consensus for a long-term R&D effort to develop a pre-competitive modeling framework for designing high functional materials. ASPRONC provided seed money for the activity of the working group and supported several international workshops organized by the working group.

In 1996, MITI approved a feasibility study for "Computational Material Design," and in 1997 a comprehensive proposal that included polymers, catalysis, parallel computation, and databases was made, but only the polymer part of the proposal was approved. In April 1998, the polymer project officially began under the leadership of Professor Masao Doi at Nagoya University.

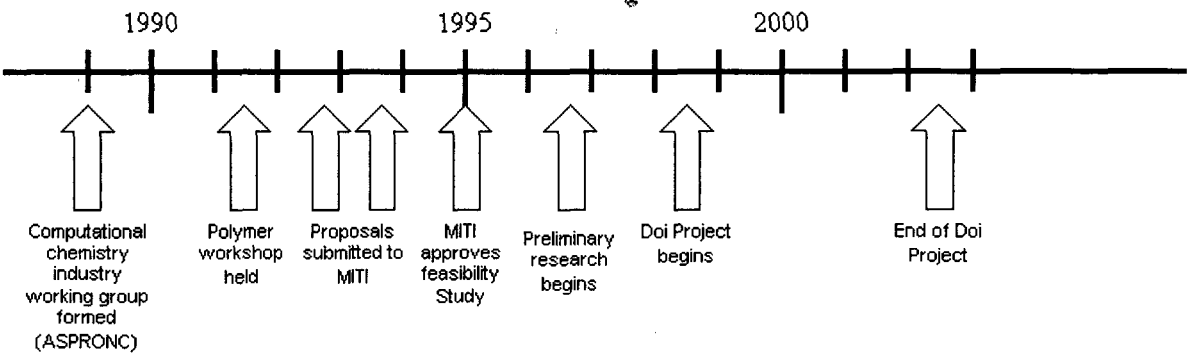


Fig. B2.1. Timeline of events in the funding and implementation of the Doi Project.

R&D ACTIVITIES

The Doi Project involves participants from twelve companies and six other university groups (Table B2.1). The university groups, who together receive 5% of the total funding, will provide experimental validation of the code developed in the Doi Project. In addition to Professor Doi, several associate professors in the Computational Science and Engineering (CSE) Department at Nagoya University are involved in the project. Students and post-docs in the CSE department participate in both fundamental research and code development. Each of the companies has assigned at least one researcher to the Doi Project full time. In all, 16 researchers are working together at Nagoya University under Professor Doi's supervision for the duration of the project. Their salaries and benefits are paid by the NEDO grant.

Reporting is done at different levels. For two to three hours a week, the on-site researchers meet and review their work. In addition, there are monthly reports to the sponsors and quarterly reviews.

Since October 1, 1999, patent rights can go to the companies involved in a particular project, as well as to Nagoya University. Originally, any patents went to NEDO, JCII, and Nagoya University. University patents would actually be owned by the university's research foundation.

Table B2.1
Company and University Participants in the Doi Project

Companies	Universities
Asahi Chemical Industry (Takeshi Aoyagi, rep.)	Yamagata University, Faculty of Engineering
JSR Co.	University of Tokyo, Institute of Industrial Science
The Japan Research Institute	University of Tokyo, Faculty of Engineering
Japan Polyolefins (Hiroyasu Tasaki, rep.)	Tokyo Institute of Technology, Faculty of Engineering
Fuji Photo Film (Hiroo Fukunaga, rep.)	Kyoto University, Faculty of Engineering
Fujitsu	Kyoto University, Institute for Chemical Research
Mitsubishi Chemicals	Nagoya University
Mitsui Chemicals (Takashi Igarashi, rep.)	
Nippon Zeon	
Sumitomo Chemical	
Toshiba Co.	
Ube Industries (Tatsuya Shoji, rep.)	

Project Goals

The goal of the Doi Project is to develop a general, object-oriented computational framework that a researcher can use to investigate the behavior and properties of multicomponent polymeric materials over a

wide range of length and time scales, from the motion of individual chains to the development of micron-sized phase-separated domains to the fracture of bulk polymers. If successful, the result will be a simulation tool that allows the user to "seamlessly zoom" between length scales in the course of a single simulation.

The team has proposed to incorporate the following simulation techniques into the framework:

- Classical explicit/united atom modeling of polymers via molecular dynamics methods
- Bead-spring models of polymers via molecular dynamics methods
- Dynamic self-consistent mean field theory
- Time-dependent Ginzburg-Landau via finite-difference methods for simulating mesoscale structure during phase separation
- Finite element modeling of constitutive equations for predicting flow and mechanical properties

The team has proposed to develop software that will simulate a wide range of phenomena, including:

- Prediction of shear and elongational viscosity and relaxation modulus for linear, branched, monodisperse and polydisperse polymer materials
- Prediction of structure, glass transition temperature, melting and crystallization temperatures, solubility parameters, elastic modulus, optical transparency and diffusion constants
- Mesoscale structure in immiscible homopolymer and copolymer blends
- Dynamical behavior of phase-separated polymers under shear
- Swelling and shrinking of polymer gels

To accomplish their goals, the team plans first to develop and validate code for each method, then to develop both the theory and techniques required to bridge the various methods in order to incorporate them seamlessly into a unified simulation framework. As a first step, Professor Doi and the members of the project have chosen low-molecular-weight polyethylene as the target material because of its industrial interest, rich phase behavior, relative simplicity from the simulation perspective, and because it is non-proprietary.

Software Development

The industrial researchers in the Doi Project carry out most of the code development in collaboration with students and post-docs at the university. Most of these researchers were trained at least to some extent in molecular or materials modeling in graduate school. They typically do not have experience in object-oriented programming, but they learn it upon joining the Doi Project.

The researchers are divided into project teams, each of which focuses on a specific simulation method. Although the industry researchers developing the software have particular applications in mind (see below for examples), their focus is on general framework development. The code is object-oriented and written in C++ using Visual Studio, a Windows-based development tool. The software is being developed for single (non-parallel) PC platforms. The user interface has been contracted out for development by Japan Research Institute, a Japanese software company. Although several MSI software tools perform functions similar to some of the tools being developed by the Doi team, the researchers do not use any MSI software products because of the unavailability of source code. Professor Doi remarked that without source code, users cannot tailor the software for their own specific needs, and often MSI software does far less that what is needed for a particular application.

At the end of the project, the software is to be distributed freely. There is a strong commitment to open-source distribution.

Application Areas

The industrial researchers in the Doi Project are interested in a wide range of applications. Two examples are Fuji Photo Film, whose goal is to develop software that will allow them to predict the phase diagram of a

liquid crystal/polyethylene blend via MD simulation, and Asahi Chemical, whose goal is to develop a general purpose MD code to predict the mechanical properties of a thermoplastic elastomer.

Intellectual Property

Issues regarding patent rights and intellectual property are still being resolved because of a recent change in the law. Under the new law, NEDO and the Japanese government claim no intellectual rights to the software developed in the Doi Project. Instead, the rights are shared by JCII and the companies participating in the project. The Doi Project management has not yet decided on its policy and plan for software distribution, such as whether the software will be freely distributed on the Internet.

Perspective of Participating Companies

During the panel's visits with several of the companies whose researchers participate in the Doi Project, the panel talked with the managers of the computational science groups in which the researchers normally work. They were all extremely eager and optimistic about the project, commenting that mesoscale modeling methods, such as those being developed in the project, will become extremely important to their R&D efforts in the very near future. They also commented that even if the Doi Project does not accomplish all of its goals, their researchers will have (a) been trained in object-oriented scientific simulation; (b) formed scientific, collaborative relationships with researchers at competitor companies, and (c) learned multiscale modeling techniques and approaches that they can apply to their company's problems. Thus, the managers believe that this is an excellent investment of both their researchers' time and their companies' resources.

CONCLUSIONS

A common theme encountered by the panel during our visit to Japan was the strong commitment by the Japanese to the development of generic, user-friendly molecular and materials computational frameworks capable of modeling a wide range of materials phenomena. Most of the code-development efforts we learned about (e.g., at NEC and JRCAT) were focused on single methods (e.g., Car-Parrinello MD), and only some involved industrial participation.

The Doi Project was thus unique in at least two respects: (1) the long-term, day-to-day collaboration of more than a dozen industrial researchers within a university setting; and (2) the ambitious goal of seamlessly integrating several complicated and disparate modeling techniques into a unified, user-friendly computational framework.

This uniqueness poses certain organizational challenges not normally faced by the molecular modeling community in Japan. Professor Doi also notes the need for new theory to underlie mesoscale modeling, expressing hope that the software tools can encourage theoretical contributions. How the team copes with these challenges will provide a valuable learning experience for future projects both in Japan and elsewhere.

Site: **Eisai Co. Ltd.**
Tsukuba Research Laboratories
Tsukuba, Japan
http://www.eisai.co.jp/
(Eisai Inc., the U.S. subsidiary, is at http://www.eisai.com/)

Date Visited: 11 November 1999

WTEC Attendees: P.A. Kollman (report author), A.M. Chaka, P.T. Cummings, P. Vashishta, H. Morishita

Hosts: Dr. Yoshiyuki Kawakami, Research Manager, Computer-Aided Drug Design
Telephone: +81-298-47-5653, Fax: +81-298-47-5771
E-mail: y1-kawakami@eisai.co.jp
Dr. Mari Ishida, R&D Planning and Coordination Division

BACKGROUND

Eisai is one of the 30 largest pharmaceutical companies in the world. Total employment exceeds 7,200 worldwide. Principal subsidiaries are located in the world's major pharmaceutical markets in the United States, Europe, and Asia. In addition to the production of pharmaceuticals, the company also produces pharmaceutical production machinery/equipment, veterinary products, bulk vitamins, and food additives.

Dr. Kawakami serves as head of the Modeling Group, which consists of four persons. From 1990-95, this group was part of the section of Structure Analysis, but since 1995, the group has been included in exploratory research. Other Modeling Group members have backgrounds in computer-assisted design (with Itai at the University of Tokyo), in organic/synthetic chemistry, and in analytical chemistry. Most have participated in the group for six to ten years.

Dr. Kawakami received his BA in Pharmaceutical Sciences from the Hiroshima University School of Medicine and his PhD in the Chemistry of Natural Products from the University of Tsukuba. Dr. Kawakami also did post-doctoral work at the University of Illinois at Chicago from 1988 to 1990 under the guidance of Professor A.J. Hopfinger. He has served as head of the Modeling Group since 1990.

Dr. Ishida is a member of the strategic research group of the R&D Planning and Coordination Division. She received her PhD in physics from the University of London.

R&D ACTIVITIES

The modeling group provides computational support for several major exploratory projects. Eisai has approximately 80 organic chemists in exploratory research, and the Modeling Group has extensive interactions in project participation. The Modeling Group focuses efforts on projects where target structures are available. Modeling is used principally in lead generation up to the nanomolar binding level and is not thought to be critical to optimization beyond that. DOCKing is a key technology in their structure-based design efforts.

The credibility of modeling in the company has been significantly enhanced by the success of Aricept (http://www.aricept.com/), an Alzheimer's disease medication for which modeling played a key role in its discovery (Sugimoto et al. 2000).

The group has several SGI Octanes and an Indigo2. The software is mainly MSI—InsightII/Discover, Cerius2, Quanta/CHARMm. They also use DOCK (UCSF) but no longer use AMBER.

In related work, there are some crystallographers and an NMR expert in the structure analysis section who work closely with modeling. There is separate work in cheminformatics and in bioinformatics.

DISCUSSION

The number of projects with known or "modelable" targets is increasing and the strategy is to participate in those projects that offer the greatest potential.

Approximately half of the organic chemists who interact with the Modeling Group have interest in modeling. Due to resource limitations, the Modeling Group unfortunately cannot meet the demands/requests of each organic chemist.

The Modeling Group maintains contact with companies specializing in modeling software development, and its members attend a variety of related scientific meetings. The group also maintains close contact with a company lab in Kyoto specializing in genomics.

The biggest challenge is perceived to be accurately treating the "free-energy term"—given that X-ray can provide structure, the critical step is to calculate the free energy of binding to 1-2 orders of magnitude. Since 90% of the computer time is spent in DOCKing and solvated MD refinement, more computer power is clearly needed. There will be a need for more modelers in the company as more targets for genomics become available. Since drug inhibitors may form reversible covalent complexes, more accurate representations of QM/MM are clearly required.

REFERENCE

Sugimoto, H., Y. Yamanishi, Y. Iimura, and Y. Kawakami. 2000. *Curr. Med. Chem.* vol. 7, pp. 303-339.

Site: **Fujitsu Ltd. and Fujitsu Laboratories**
Computational Science & Engineering Center
9-3, Nakase 1-chome, Mihama-ku
Chiba 261-8588, Japan

Date Visited: 12 November 1999

WTEC Attendees: A.M. Chaka (report author), R.S. Jones, P.A. Kollman, H. Morishita, P. Vashishta

Hosts:

Fujitsu Ltd.

Munetaka Takeuchi, Computer Chemistry Systems Dept., Computational Science & Engineering Center, Fujitsu Limited
Tel: (81) 43-299-3236, Fax: (81) 43-299-3019
take@strad.se.fujitsu.co.jp

Masahito Yamaguchi, Manager, CCS Marketing Sec., Computer Chemistry Systems Dept., Computational Science and Engineering Center, Fujitsu Limited
Tel: (81) 43-299-3236, Fax: (81) 43-299-3019
yamaguch@ssd.se.fujitsu.co.jp

Takashi Mitsui, Computational Physics Application Section, Computational Science Application Laboratory, Research Center for Computational Science, Fujitsu Limited Tel: (81+43)299-3240, Fax: (81 +43)299-3010

Fujitsu Laboratories Ltd.

Toshiaki Narusawa, General Manager, Materials and Material Engineering Laboratories, Fujitsu Laboratories Ltd., 10-1, Morinosato-wakamiya, Atsugi 243-01, Japan
Tel: (0462)-50-8828, Fax: (0462)-48-8812

Azuma Matsuura, Senior Researcher, Functional Organic Materials Laboratory
Tel: 81-462-50-8257, Fax: 81-462-48-6000
E-mail: az@flab.fujitsu.co.jp

BACKGROUND

The panel met with representatives from two Fujitsu subsidiaries, Fujitsu Limited (describing commercial software development) and the Materials and Material Engineering Laboratories of Fujitsu Laboratories (describing modeling applications). Hence, the discussion is divided into two parts, one for each subsidiary.

The consolidated Fujitsu Group is a $43.3 billion company with 188,000 employees, 132 subsidiaries in Japan, and 386 subsidiaries overseas. As a percentage of sales, 39% comes from software and services, 34% from information processing hardware, 13% from telecommunications, 10% from electronic devices, and 4% from other product categories. A wholly owned subsidiary, the NIFTY Corporation, is the top internet service provider in Japan with 350 million users worldwide. Computers manufactured by Fujitsu range from Intel architecture servers to Unix servers to global servers (mainframes).

In 1997 Fujitsu established "Green Product" standards for its products, which are based on Germany's "Blue Angel" eco-mark. Their goal is eventually to manufacture all of their products to meet these demanding environmental standards.

FUJITSU LTD.

The software and services business resides in Fujitsu Ltd., a wholly owned subsidiary with 55,000 employees, which ranks second worldwide in sales, at $16 billion annually. This compares to IBM, the premier international software and services provider, at $40 billion in annual sales. Fujitsu's slogan is "Everything on the Internet."

In deciding whether or not to purchase a supercomputer, the quality and quantity of software available is becoming increasingly as important as high performance. Fujitsu has adopted two approaches to ensure the availability of software for its computers: to port *de facto* standard application packages, and to develop software internally in fields that are new and rapidly expanding. Application packages are developed by collaboration with leading customers and researchers, by participation in national development projects, or by direct commercialization of Fujitsu-originated technologies.

Typical application packages for VX/VPP computers include software for collision and impact analysis, structure and vibration analysis, computational fluid dynamics, electromagnetic field analysis, computational chemistry, and mathematical libraries.

The goal of the Computer Chemistry Systems Department in Fujitsu Limited is to develop and sell computational chemistry and physics software. There are two Fujitsu-exclusive products for the advanced computational scientist, MOPAC2000 and MASPHYC. MOPAC 2000 is a patented product developed by Dr. James Stewart and Fujitsu scientists with the following features:

- Mozyme—Order(N) NDO semi-empirical quantum mechanical program designed for large systems such as enzymes
- MOS-F—configuration interaction using INDO/S and CNDO/S for calculating molecular spectroscopy
- Quantum dynamics

MASPHYC and MASPHYC SP (surface phenomena) stand for "material design system by means of computational chemistry and physics." The purpose of the software is to predict phase behavior and transformations of organic, ceramic, and metallic systems by using molecular dynamics (MD). It includes database management and properties analysis, such as transport coefficients, phonon dispersion spectrum, and elastic constants. The current implementation of MASPHYC is based on classical MD. The expandable potential library currently contains 2-body potentials (Lennard-Jones, BMH (Gilbert-Ida), Morse, and Johnson), 3-body potentials (Stillinger-Weber, Tersoff), and several Embedded Atom (EAM) methods. Fujitsu is looking into the possibility of incorporating quantum dynamics into the MASPHYC framework with such programs as FHI-96MD (from the Fritz-Haber-Institut in Berlin, Germany), VASP (U. Vienna), and WIEN-98 (FP-LAPW from T.U. Vienna).

Examples of successful MASPHYC applications are as follows:

- Pressure-induced phase transition of $AlPO_4$
- Simulation of the SmC phase of ferrelectric liquid crystalline molecules (Toriumi et al. 1996)
- Adsorption of methane on silicalite
- Thin-film growth of hydrogenated amorphous silicon
- Vacancy migration in BaFCl (Kurobori et al. 1999)
- Searching possible conformations in benzene molecular clusters (Li et al. 1995)
- Effects of protein electrostatic environment on the absorption maximum of bacteriorhodopsin (Hojou et al. 1998)

Fujitsu also has software for the beginning computational or bench chemist, WinMOPAC, which contains MOS-F, MOPAC 97, and QSAR/QSPR. In addition, there is ChemFrontier, a chemical structure drawing tool, which does 2D to 3D structure interconversion and can cut and paste structures to and from Microsoft Office.

[In March 2000, Fujitsu Limited announced acquisition from Oxford Molecular of the CAChe product line of molecular-modeling software for PC and Macintosh computers (*Chem. & Eng. News,* March 27, 2000; http://www.oxmol.com/news/fujitsu.shtml). This product line would remain headquartered in Beaverton, Oregon. Efforts toward a smooth transition will be aided by nine-month joint-development agreement between the companies. One announced modification will be to replace CAChe's existing MOPAC software with the more powerful MOPAC 2000 from Fujitsu.]

FUJITSU LABORATORIES LTD.

Molecular and Materials Modeling Activities

In 1962 Fujitsu merged separate divisional R&D efforts into one central "company" called Fujitsu Laboratories. Today, Fujitsu Laboratories has seven labs and two centers, one of which is in the United States. There are several project groups whose purpose is to work across boundary lines and facilitate communication between the various centers and groups. Fujitsu Laboratories is a ¥5 billion ($50 million) company with 1,500 employees. The ratio of R&D expenditures to net sales for Fujitsu as a whole has declined over the past five years, from 9.9% in 1995 to 7.5% in 1999.

Its research activities are extremely broad and diverse, but the focus is to "develop core multimedia technology." Fujitsu Laboratories is divided into Multimedia Systems, Personal Systems, Network Systems, Electronic Devices and Materials, Materials & Material Engineering, System LSI, Project Groups and Center, plus Fujitsu Laboratories of America, Inc. Major research efforts focus on shifting from DRAMs to system-on-a-chip architecture, flash memory, compound semiconductors, and plasma display panels (PDP) to improve the speed, mobility, and multimedia capability of systems on the Internet.

At Fujitsu Laboratories, the expressed purpose of molecular and materials modeling using computational chemistry is to find appropriate materials efficiently, particularly electronic and optical materials. Modeling researchers are in all research divisions, typically with 3-4 modelers in a group. The research is very applied, and modeling is generally used to screen candidates, interpret experimental results, and to predict properties that cannot be obtained with experiments.

Molecular and materials modeling successes were highlighted in several areas of research, including biological information science and the human genome project, optimizing the Si-SiO_2 interface structure for the design of CMOS devices, and chemical vapor deposition.

The Functional Organic Materials Laboratory was chosen as a representative area for the panel's visit, and it exemplifies how molecular and materials modeling is integrated in the R&D process. (The Personal Systems Lab is also a big user of modeling.) In the Functional Organic Materials Laboratory, researchers developed a semi-empirical quantum mechanical method based on neglect of differential overlap (NDO), similar to ZINDO, to predict hyperpolarizability and optical spectra. In the DFT area, several people are collaborating with Prof. Terakura at JRCAT.

One of the biggest successes has been in developing a resist material for microlithography. Fujitsu needed a short-wavelength absorber, and modeling was used to calculate absorption wavelengths to screen candidates and focus development direction.

Fujitsu researchers use Gaussian 98, MOPAC2000, and WinMOPAC most heavily. Internally they have developed a ZINDO-based program called MOS-F in order to predict the spectroscopic properties of large organic systems. The latest version is MOS-F V4.2c, about 30,000 lines of code long. They also have developed a UNIX X-windows viewer of molecules, wavefunctions, and electron density.

DISCUSSION

Dr. Matsuura emphasized that the researchers at Fujitsu Laboratories find it very important to predict physical and chemical properties of molecules with moderate size quantitatively and with large size qualitatively. They have applied molecular orbital calculations to predict electronic and optical properties and chemical reactivities. However, their calculation was limited to molecular systems. They expect that current modeling methods that manipulate electron behavior will be extended to mesoscopic systems or materials that have no periodic boundary condition. Thus, the possibility of creating many novel functional materials will increase.

Researchers expect that increased computer power and new theories and methods will have favorable impacts. Improvements in computer technology will yield enhanced processing speed and greater memory

size in line with existing trend curves. Parallelization is one of the best ways to speed up the computer. Even the free UNIX operating systems such as Linux and FreeBSD are compatible with parallel computer systems, and personal computers have dual CPUs. The ordinary personal computer will be 64 bits and will have multiple CPUs in 5 years, reaching performance over 1 Gflop. Mid-range and high-end computers will manipulate over 1 terabyte of memory and their performance will be over 1 Tflop. If so, they imagine calculating the electronic structure of mesoscopic systems qualitatively, along with real materials properties. Twenty years ago, thousands of people shared a very slow computer system. Dr. Matsuura recalled using the computer system with maximum memory of 3 MB at the computer center of University of Tokyo, which was the largest computer center in Japan at that time. Now, the Pentium/Linux systems with 256 MB RAM are cheap enough that people can buy them individually.

Nevertheless, these enhancements may not be enough for the applications of molecular modeling using computational chemistry. There are two reasons for this consideration. One reason is the need for accuracy. Another is that materials in the field of electrical devices can require an astronomical size of calculations, because they include many types of phenomena or properties of molecules and atoms, which are very complicated and have huge size. Fujitsu's MOPAC2000 is currently an effective tool for calculations on huge materials systems.

Accordingly, in theory and algorithms, they think the present breakthroughs are the linear scaling method and the spatially hybrid ONIOM method of Morokuma. In the past 15 years, direct methods such as direct SCF method and efficient evaluation of two-electron integrals, such as Obara-Saika and Prism algorithms, were developed and became widespread. Such methods have an important role in molecular modeling. However, problems also occur. For example, they have usually encountered SCF-convergence failure when calculating an anion molecule or excited states of a molecule with diffuse functions. Also, computational cost increases radically with the size of system using conventional MO theory. ONIOM and Order(N) methods provide some valuable alternative approaches, as should the seamless zooming model of Professor Doi.

Testing and validation are important. The decision to do round-robin tests or not depends on whether there are molecular properties to predict. For example, calculated optimized geometry of organic molecules at the HF/6-31G* or higher level of theory can be compared among the different groups of organic molecules. In contrast, calculated excitation energies of organic dyes cannot be compared. Reliability is established mainly by comparison with experimental data. Methods are then chosen by a mixture of experience, inspiration, and trial and error.

Fujitsu managers see cooperation between experimental researchers and theoretical ones as most important of all. Without the cooperative work, theory would not move beyond the theory itself and into materials development. Regardless of the advances in computer technology in the next 15 years, it will not be enough to satisfy the highest demands of researchers in the theoretical field. At the same time, for experimentalists it is very important that the software work well on Windows 95/98 machines with Windows-style GUIs. Special GUIs or command-line input prevents such people from using it.

Dr. Matsuura stated that another important aspect is the further globalization of software development. Especially, academic research and development of calculation theories and programs should be achieved and applied with international co-operation. One reason he thinks so is that international co-operation on the development of calculation programs meets the expectation of humankind since it helps the development of science and technology. Another reason is that superior researchers and highly skilled programmers are very rare, even worldwide.

Finally, they think that it is very important to construct dedicated computer systems. In the first stage of material or device development, the calculation is performed roughly on small-size computers such as PCs for the screening of raw materials. In the next stage, *ab initio* MO calculations with high accuracy are performed on high-end computers to be compatible with the experiment. In these circumstances, data linkage between the high-end computer and small computer might be important for the realization of high-performance molecular modeling. In addition to this linkage, linking to the Internet and global databases might be important for higher-level molecular design for green chemistry.

Government-Sponsored Projects

Fujitsu participates in government-sponsored projects such as the Doi Project (see separate report). These projects typically begin with two years of government support for exploratory research. After two years, a five-year plan will be developed with more extensive funding and more focus.

In the past, there was a strong distinction between projects. For example, computer centers founded for one purpose could not be used for another. This is gradually changing.

With MITI-sponsored projects, there are two ways to proceed. In one, many people from industry go to national labs and form a virtual research lab. The other way is to divide the research project into several separate targets that researchers then do within their own companies. Usually MITI projects are done by industry and academic participation is rare, but this is also changing. This was strongly influenced by the United States.

CONCLUSIONS

Fujitsu Ltd. is endeavoring to sell a wide range of computational chemistry and physics software for a broad range of users on platforms ranging from Windows/NT to large massively parallel vector machines. They have exclusive capabilities with MOPAC2000 and MASPHYC in the commercial molecular and materials software arena.

In Fujitsu Laboratories, molecular and materials modeling is seen as an important part of their R&D efforts and is integrated into every research department.

REFERENCES

Hojou, H., et al. 1998. *Chem. Phys. Lett.* 294, 162.

Kurobori, T., et al. 1999. *Rapid Research Notes, Phys. Stat. Sol (b)* 212, R5.

Li, Z., et al. 1995. *Comp. Mat. Sci.* 4, 241.

Toriumi, H., et al. 1996. *J. Phys. Chem.* 100, 15207.

Site: **Institute for Molecular Science (IMS)**
Okazaki, Aichi 444-8585
Japan
http://www.ims.ac.jp

Date Visited: 10 November 1999

WTEC Attendees: K. Morokuma (report author), P.R. Westmoreland, P.T. Cummings, S.C. Glotzer, R. Miranda, E.B. Stechel

Hosts: Dr. Suehiro Iwata, Prof., Department of Theoretical Studies, and Director, Computer Center, iwata@ims.ac.jp
Dr. Fumio Hirata, Prof., Department of Theoretical Studies, hirata@ims.ac.jp
Dr. Yuko Okamoto, Assoc. Prof., Department of Theoretical Studies, okamotoy@ims.ac.jp

BACKGROUND

The Institute for Molecular Science (IMS, http://www.ims.ac.jp), established in 1975, is one of about ten national "inter-university" institutes in Japan. Uniquely in the same campus, there are two other inter-university institutes, National Institute for Basic Biology (NIBB, http://www.nibb.ac.jp) and National Institute for Physiological Science (NIPS, http://www.nips.ac.jp), which with IMS constitute the Okazaki National Research Institutes (ONRI, http://www.orion.ac.jp) with a common administration bureau.

The mission of IMS is two-fold: to be the center of excellence in its own research and to provide services to the university community. It has 17 professors and 30 associate professors (in 47 research groups), along with 50 research associates (junior faculty) and 42 technical associates. IMS is also a part of the rather new Graduate University for Advanced Studies (http://www.soken.ac.jp) and has 48 graduate students (all PhD students with MS degrees) of its own and 28 from other universities. It also has approximately 80 postdocs and visiting scientists (about 40 domestic and 40 foreign). Its staff members published over 400 scientific papers in 1998. The annual budget is about $40 million (of which $12 million is for personnel). In addition, the faculty brings in about $2.5 million in research grants/contracts from the government and $1 million from industry. Among the service centers are the Computer Center, Ultraviolet Synchrotron Orbit Radiation Facility, Laser Center, and Research Center for Molecular Materials. Research activities of the IMS are summarized in IMS's Annual Review (http://www.ims.ac.jp/publications/), which is published usually in December each year.

COMPUTER CENTER

The computer center (http://ccinfo.ims.ac.jp), established in 1977, is a service center of IMS. It is equipped with powerful supercomputer systems and provides a large amount of computer time to the molecular science community throughout Japan. It had a major impact on the development of molecular modeling computation at IMS and Japanese universities in the 1980s, when local computational resources were extremely limited, and it still provides excellent computer CPU and software services to the molecular science community. Until 1995, it provided substantial computational support to the solid-state physics community, prior to the establishment of the supercomputer center at the Institute for Solid State Physics. It presently leases as major computer systems NEC SX-5 (8 vector processors with 32 GB shared memory), NEC SX-3 (3 vector processors with 2GB shared memory), NEC HPC (2 vector processors with 1GB shared memory) and IBM SP2 (48 nodes) with the total annual rental fee of about $7 million. The SX-3 is an old system and is scheduled for replacement in March 2000 by a system that is about 20 times faster than the present one.

The center includes a director (Prof. Suehiro Iwata), associate prof. (Mutsumi Aoyagi), three research associates, and five technical staff members. Presently about 70% of the computer time is used by outside users and 30% by IMS users. There are approximately 170 outside projects (with 600 outside users), which were reviewed and approved by the executive committee of the center, consisting of four IMS, one NIBB,

one NIPS and four outside members. The center supports many commercial and academic codes in molecular science. The center also sponsors development of the Quantum Chemistry Literature Database (QCLDB), a database of *ab initio* and density functional calculations since 1978, and makes it available via its Web site (http://qcldb.ims.ac.jp) and tape distribution. No industrial users are allowed to access the center's facilities.

RESEARCH ACTIVITIES IN THEORETICAL/COMPUTATIONAL MOLECULAR SCIENCE

Computational molecular science activities at IMS are mainly carried out at the Department of Theoretical Studies (DTS) and the Computer Center. Extensive collaborations take place between these theoretical groups and experimental groups inside as well as outside the IMS (domestic and overseas). Some experimental groups at IMS also perform some modeling work in connection with their own experiments.

The Department of Theoretical Studies consists of three professors, three associate professors, nine research/technical associates, approximately 20 postdocs/visiting scientists and approximately 15 graduate students. An associate professor, three research/technical associates and a postdoc at the Computer Center are also involved in computational/theoretical research. The theoretical groups, as well as other groups in IMS, are very international, with several international visiting fellows and postdocs. Although presently there is no visiting fellow from industry in the computational groups at IMS, in the past several people from industry spent one to two years at IMS for training and collaboration. Several former graduate students and postdocs at IMS also went into industrial laboratories. These former IMS associates are playing rather important roles in industrial applications of molecular modeling in Japan.

The faculty members involved mainly in modeling activities are as follows:

- Dr. Suehiro Iwata, Professor, DTS, quantum chemistry, spectroscopy, solvation clusters (http://zeus1.ims.ac.jp/Iwata/, iwata@ims.ac.jp)
- Dr. Yuko Okamoto, Associate Professor, DTS, protein folding (http://konf2.ims.ac.jp, okamotoy@ims.ac.jp)
- Dr. Hiroki Nakamura, Professor and Director, DTS, quantum and semi-classical molecular collision and reaction dynamics (http://www.ims.ac.jp/e/nakamura.html, nakamura@ims.ac.jp)
- Dr. Yoshitaka Tanimura, Associate Professor, DTS, electronic properties and spectroscopy in condensed phase (http://fuji.ims.ac.jp, tanimura@ims.ac.jp)
- Dr. Fumio Hirata, Professor, DTS, solution, solvation, liquid-solid interface (http://daisy.ims.ac.jp, hirata@ims.ac.jp)
- Dr. Kenji Yonemitsu, Associate Professor, DTS, solid-state physics, organic conductors (http://magellan.ims.ac.jp, kxy@ims.ac.jp)
- Dr. Mutsumi Aoyagi, Associate Professor, Computer Center, potential energy surfaces and dynamics of small excited-state molecules (aoyagi@ims.ac.jp)

The details of their research activities can be found in the IMS Annual Review mentioned above, as well as on IMS's Web sites.

One of the recent foci in the activities of these groups is the recent major research grant on protein structures. This grant of $4 million over five years (1998-2003) is entitled "Development of simulation algorithms for the first-principles prediction of three-dimensional structures of proteins" (http://konf2.ims.ac.jp/okamotoy/mirai.html). It is supported by the Japan Society for the Promotion of Science (JSPS, http://www.jsps.go.jp) Research for the Future (RFTF) Program under the research promotion field of "Computational Science and Engineering (CSE)" (http://olab.is.s.u-tokyo.ac.jp/cse). The JSPS RFTF program, started in 1996 and funded by the Japanese government, is intended to promote and expand the frontiers of scientific research. Seven projects have been funded in the CSE field.

The principal investigator (PI) of this grant is Prof. Okamoto, with Prof. Hirata and Prof. Aoyagi as co-PIs. This project involves some collaboration with universities and NIPS, but not with industry. Prof. Okamoto summarized the goal of the present project as follows:

There is a close relationship between the three-dimensional structures of proteins and their biological functions. The study of protein structures is thus aimed at the understanding of how proteins carry out their functions. The research in this field is ultimately led not only to drug design and de novo design of artificial proteins with specific functions but also to the elucidation of the pathogenic mechanism for the illness that is caused by misfolding of proteins (such as mad cow disease and Alzheimer's disease).

It is widely believed that the three-dimensional structures of proteins are determined solely by their amino-acid sequence information. However, the prediction of protein structures by computer simulations with the input of only the amino-acid sequence (prediction from the first principles) has yet to be accomplished. There are two elements of the difficulty. One is that the number of internal degrees of freedom of protein systems is extremely large, and there exist a huge number of local minima in the energy function. It is a very challenging problem to find the global-minimum state in free energy, which corresponds to the native protein structure, since simulations by conventional algorithms will get trapped in one of the local-minimum states. The other element for the difficulty is the problem of accuracy in the energy function of the protein system including the solvent. It is obvious that if the energy function is crude, then one cannot expect to have the right prediction. Proteins *in vivo* are surrounded by a solvent with water and ions, but water is one of the most complicated liquids. It is also non-trivial to calculate the electronic structures of a large system like proteins.

The goal of the present project is to succeed in the prediction of the three-dimensional structures of proteins from the first principles by overcoming the above-mentioned difficulties (namely, by developing a simulation algorithm that can avoid getting trapped in states of energy local minima and a numerical method that can evaluate the energy function both accurately and efficiently). In particular, we try to predict, for the first time, the three-dimensional structure of a small protein with about 50 amino acids.

Prof. Okamoto discussed two important problems to be addressed in obtaining the native 3D structure of a protein that corresponds to the state of the global minimum in free energy. The first problem is the use of a more precise energy function than before. They used ECCP/2 type energy functions consisting of electrostatic + Lennard-Jones + hydrogen bond + torsion terms. The effect of the solvent was included as the dielectric constant (level 1), by the solvent accessible surface area (level 2), or with more realistic models, such as scaled particle theory, RISM and explicit water molecules (level 3). Based on preliminary calculations for small peptides, the following future subjects were pointed out: the use of larger peptides, test of energy functions and parameters other than ECCEP/2 type, and examination of the salt effects. The second problem is to use better algorithms to avoid being trapped in local minima. The multicanonical algorithm by Berg and Neuhaus has been tested by his group and has turned out to be a very efficient method for optimization. He recently found that the replica-exchange method (REM, also referred to as parallel tempering) (Hukushima and Nemoto 1996) is extremely well suited for protein-folding simulations. The REM-MD method has been developed and is being applied to the structure of small polypeptides.

Professor Hirata emphasized the merit of his quantum-classical hybrid approach in solution chemistry, in particular the combination of his XRISM method (Hirata and Rossky 1981) with various other approaches. The XRISM-SCF method combines the RISM with *ab initio* MO methods. It can treat the electronic structure in solution and is suited for chemical reactions in solution. The strong points of this method are (1) solvent in molecular detail, (2) proper thermodynamic limit, (3) reasonable computer load (that can be improved), and (4) no adjustable parameters. The weak points are (1) more difficult to learn background, and (2) some mathematical approximations based on the dielectric constant, although that can be improved. The XRISM+MC simulation is suitable for protein folding. The XRISM+GLE (generalized Langevin equation) method can handle non-equilibrium processes in liquid. The XRISM+DFT method has been used for study of the electrode-solution interface. For a recent review of the approaches, see Hirata (1998).

Professor Iwata summarized his research on electronic structure of molecular systems using two examples. As the first example, he showed the results of accurate *ab initio* calculations of the potential energy surface of the ground electronic state of the CO molecule, and the detailed comparison with recent high-resolution

spectroscopy experiments. As the second example, he presented results of *ab initio* MO calculations for $M(H_2O)_n$ cluster, where M is an alkali metal. With n=5 and M=Li, he recently found a new structure, in which a cluster of water molecules is formed around the lithium cation and the excess electron is trapped between two water molecules in the cluster by forming a bond that may be called "the electron hydrogen bond." The emergence of this structure is consistent with the experimental flattening of the ionization potential for n>5. The first example is a part of his new three-year research project on "Computational Chemistry for Molecular Spectroscopies and Chemical Reactions in Atmospheric Environmental Molecules," which is funded by the new Research and Development Applying Advanced Computational Science and Technology program of the Japan Science and Technology Corporation (JST, a corporation under the Science and Technology Agency).

CONCLUSIONS

IMS is the most important center in molecular modeling in Japan (and certainly one of the best in the world), and has a very high concentration of active researchers (~50 people) in the field. It covers a very wide range of interests, including quantum chemistry, quantum and semi-classical collision dynamics, modeling of solution and solvent effects, modeling of protein folding, electronic properties in condensed phase and solid state physics. Although direct connections to industry are not strong, it has provided personnel and training for industrial molecular modeling and also has played a leading role in computational molecular science for both academics and industry. The new JSPS protein-folding project is very ambitious, but, if successful, it will have a large impact on progress in protein modeling.

REFERENCES

Baeck, K.K., H. Choi, and S. Iwata. 1999. Theoretical study on spectroscopic properties of positive, neutral and negative species of BCl2 and AlCl2: The stability of the negative species. *J. Phys. Chem. A* 103, 6772-6777.

Bandyopadhyay, P., S. Ten-no, and S. Iwata. 1999. Structures and photoelectron spectroscopies of Si2C2- studied with ab initio multicanonical Monte Carlo simulation. *J. Phys. Chem. A* 103, 6442-6447.

Fuke, K., K. Hashimoto, and S. Iwata. 1999. Structures, spectroscopies and reactions of atomic ions with water clusters," *Adv. Chem. Phys.*, 110, 431.

Hansmann, U.H.E., and Y. Okamoto. 1993. Prediction of peptide conformation by multicanonical algorithm: new approach to the multiple-minima problem. *J. Comput. Chem.* 14, 1333.

Hansmann, U.H.E., and Y. Okamoto. 1999. Effects of side-chain charges on alpha-helix stability in C-peptide of ribonuclease A studied by multicanonical algorithm. *J. Phys. Chem. B* 103, 1595.

Hansmann, U.H.E., and Y. Okamoto. 1999. New Monte Carlo algorithms for protein folding. *Curr. Opin. Struct. Biol.* 9, 177.

Hansmann, U.H.E., and Y. Okamoto. 1999. The generalized-ensemble approach for protein folding simulations. In *Ann. Rev. Comp. Phys. VI*, ed. by D. Sauffer (Singapore, World Scientific) p. 129.

Hansmann, U.H.E., M. Mausya, and Y. Okamoto. 1997. Characteristic temperatures of folding of a small peptide. *Proc. Natl. Acad. Sci. USA* 94, 10652.

Hirata, F. 1998. *Bull. Chem. Soc. Jpn* 71, 1483.

Hukushima, K., and K. Nemoto. 1996. *J. Phys. Soc. Jpn.* 65, 1604.

Ikegami, T., and S. Iwata. 1999. Photodissociation dynamics of argon cluster ions. *J. Chem. Phys.* 110, 8492-8500.

Kawai, H., T. Kikuchi, and Y. Okamoto. 1989. A prediction of tertiary structures of peptide by the Monte Carlo simulated annealing method. *Protein Eng.* 3, 85.

Kinoshita, M., Y. Okamoto, and F. Hirata. 1998. First-principle determination of peptide conformations in solvents: combination of Monte Carlo simulated annealing and RISM theory. *J. Am. Chem. Soc.* 120, 1855.

Okamoto, Y. 1998. Protein folding problem as studied by new simulation algorithms. *Recent Res. Devel. in Pure & Applied Chem.* 2, 1.

Okamoto, Y., and U.H.E. Hansmann. 1995. Thermodynamics of helix-coil transitions studied by multicanonical algorithms. *J. Phys. Chem.* 99, 11276.

Okamoto, Y., M. Fukugita, T. Nakazawa, and H. Kawai. 1991. Alpha-helix folding by Monte Carlo simulated annealing in isolated C-peptide of ribonuclease A. *Protein Eng.* 4, 639.

Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314, 141.

Ten-no, S., S. Iwata, S. Pal and D. Mukherjee. 1999. Generalization of the coupled-cluster response theory to multireference expansion spaces: an application of the coupled-cluster singles and doubles effective Hamiltonian. *Theor. Chem. Acc.* 102, 252.

Tsurusawa, T., and S. Iwata. 1999. Theoretical studies of structures and ionization threshold energies of water cluster complexes with a group 1 metal, M(H2O)n (M=Li and Na). *J. Phys. Chem. A* 103, 6134-6141.

Watanabe, N., S. Ten-no, S. Pal, S. Iwata, and Y. Udagawa. 1999. Size-extensive calculations of static structure factors from the coupled cluster singles and doubles model. *J. Chem. Phys.* 111, 827-832.

Site: **Institute for Solid State Physics (ISSP)**
University of Tokyo
Visited: 7-22-1 Roppongi, Minato-ku; Tokyo 106-8666, Japan
New location: 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8581, Japan
http://www.issp.u-tokyo.ac.jp/index_e.html

Date visited: 9 November 1999

WTEC Attendees: P. Vashishta (report author), A.M. Chaka, R.S. Jones, P.A. Kollman, H. Morishita

Hosts:
Professor Hidetoshi Fukuyama, Director, ISSP, (81)-471-36-3250, fax (81)-471-36-3251; fukuyama@issp.u-tokyo.ac.jp
Professor Hajime Takayama, ISSP, takayama@issp.u-tokyo.ac.jp
Professor Shinji Tsuneyuki, ISSP, stsune@issp.u-tokyo.ac.jp
Professor Masatoshi Imada, ISSP, imada@issp.u-tokyo.ac.jp
Professor Susumu Saito, Tokyo Institute of Technology, Department of Physics, 2-12-1 Oh-okayama, Meguro-ku, Tokyo 152-8551, saito@stat.phys.titech.ac.jp
Professor Kazuyuki Watanabe, Science University of Tokyo, Department of Physics, Faculty of Science, 1-3 Kagurazaka, Shinjuku; Tokyo 162-8601 kazuyuki@rs.kagu.sut.ac.jp
Dr. H. Akagi, University of Tokyo, Department of Physics, Faculty of Science, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033

BACKGROUND

The ISSP is the premier research institute in Japan for solid-state/condensed-matter physics research. The institute is associated with the University of Tokyo, but in fact it is a self-contained research organization.

The WTEC group was welcomed by Professor Fukuyama, Director of ISSP. Professor Hajime Takayama conducted the meeting and was the contact person for the visit. He described the background of ISSP and the organization of the supercomputer center. Scientific presentations were made by Professors Shinji Tsuneyuki, Susumu Saito, Kazuyuki Watanabe, and Dr. H. Akagi (in place of Professor Masaru Tsukada, who was unable to attend the meeting).

The ISSP was established in 1957 to carry out research in solid-state physics and provide experimental facilities for other Japanese researchers in a collaborative mode. ISSP is located on prime land in Tokyo in Roppongi, the embassy area with many restaurants and clubs. The institute is being relocated outside the Tokyo area like several other parts of the University of Tokyo. [Note: ISSP moved to its new location on April 1, 2000].

The first large-scale scientific computations for research were carried out in 1980 at university supercomputer centers. All the seven major national universities (Tokyo, Kyoto, Osaka, Nagoya, Sendai, Kyushu, and Hokkaido) established university supercomputer centers in the 1980s. The current supercomputer system at ISSP was established in 1995, and the ISSP was reorganized in 1996.

The Materials Design and Characterization Laboratory (MDCL) was established in May 1996 as the third research facility of the ISSP. The aim of MDCL is to promote materials science with an emphasis on the design, synthesis, and characterization (DSC) cycle. DSC is the integration of three functions to exploit new materials. MDCL consists of two sections: a materials design (MD) section and a materials synthesis and characterization (MSC) section. The supercomputer center of ISSP (SCC-ISSP) is in the MD section, while in the MSC section there are six laboratories for joint use—materials synthesis, chemical analysis, X-ray diffraction, electron microscope, electromagnetic measurement, and spectroscopy laboratories.

SUPERCOMPUTING SYSTEM

The SCC-ISSP contains a vector-parallel supercomputer Fujitsu VPP500/40, a file server DEC7000/640 AXP, a graphics machine SGI Onyx RE2, and many desktop workstations in the terminal room and distributed in the offices and laboratories. The supercomputer is a distributed memory MIMD system with 40 vector processors with a peak speed of 64 Gflops and has 13 Gbytes of total memory. Hardware and software and maintenance costs of the system are $0.6 million per month—$7.2 million per year. SCC-ISSP is operated by a staff of seven. Professor Shinji Tsuneyuki supervises six research and technical assistants.

The supercomputer system is placed in the service of general researchers of solid state/condensed matter physics in Japan through the user program. The SCC-ISSP has a steering committee (19 members—9 from ISSP and 10 from universities) and supercomputer advisory committee (39 members—10 from ISSP and 29 from universities). Projects can be submitted to SCC-ISSP by any researchers from Japanese universities or public research institutes. Researchers from industrial research laboratories are not allowed to compute at SCC-ISSP.

The SCC-ISSP supports computational activities for about 200 research groups located at more than 60 Japanese universities. Each group on the average consists of two to four investigators—professor, research assistants (equivalent to assistant professor in the United States), and graduate students. Its aims are to preferentially support large-scale computations, which cannot be performed at university supercomputer centers and to support the development of new computational algorithms. Computer usage by project type and change over time are summarized in Tables B2.2 and B2.3 below. (In 1998 the total service time decreased because the system was accidentally shut down many times, mainly due to troubles in the cooling equipment.)

Table B2.2
Number of Projects, Total Number on Node-Hours Requested and Approved in 1998

Class	Max Node-hours	Application	Number of Projects	Total node-hours (requested) approved
A	200	Any time	7	(1400) 1400
B	1500	Twice a year	131	(160,000) 101,000
C	none	Twice a year	28	(135,000) 100,000
D	none	Any time	3	(7000) 6000 (Requests considered urgently)
E	none	Twice a year	0	(Collaborative research by several investigators from different institutions.)

Table B2.3
Computer Use: 1995, 1996, 1997, and 1998

Year	1995	1996	1997	1998
Total service time (hours)	317,080	307,800	298,880	282,338
Number of executed jobs	56,054	67,126	51,496	30,914
Total CPU time (hours)	147,885	211,184	230,094	182,624
Availability (%)	94.4	87.8	85.3	80.6
Utilization rate (%)	46.6	68.6	77.0	64.7

The number of active users was 220 in 1998, which is about 40% of the number of registered users. Among these active users, the 50 heaviest users consumed about 67% of total CPU time. The number of active "parallel computing" users was 54, which is about 25% of the number of active users. About 67% of the total CPU time in the parallel job class was consumed by the 10 heaviest "parallel computing" users.

The SCC-ISSP's current system will be replaced in March 2000 with two new supercomputers, fileserver and network switch. This new system will include:

- Hitachi SR8000/60 60-processor vector-parallel supercomputer
- SGI Origin 2000/384 384-processor shared-virtual-memory supercomputer
- Compaq GS160/8 8-CPU fileserver

The new supercomputer system is expected to serve three roles:

1. Further enhancement of the preferential support of large-scale computing and new algorithm development projects.
2. Enhanced level of service and support to the computational condensed matter/solid state physics community. This is expected to help raise the level of computational research in the field.
3. Contribution to international cooperative research. It will be possible for foreign researchers to use the SCC-ISSP facility in collaborative projects with university-based Japanese researchers.

SCC-ISSP does not use any commercial codes for simulations on the supercomputers. All computer codes used by researchers are their own or developed with other university researchers. However, Gaussian is used on workstations that are linked to the supercomputer system.

R&D ACTIVITIES

Fields covered by SCC-ISSP can be divided into three major areas of research. Percent of total computational resources used in each area in 1998 is given in parentheses:

- First-principles calculation of material properties (45%)
- Strongly correlated electron systems (45%)
- Cooperative phenomena in complex, macroscopic systems (10%)

First-Principles Calculation of Material Properties

Users in this area of research are distributed at all major universities. The groups use all the methods of solid state/condensed matter physics. To enumerate the methods used in computations by this group of researchers, the following abbreviations are used:

- DFT: density functional theory
- PW: plane wave
- FLAPW: full-potential linearized augmented plane wave
- LMTO: linear muffin tin orbital
- LSDA: local spin density approximation
- GW: GW approximation
- NM: new method

Some of the research projects, the name of the investigator, and the computational method used are given below:

- TSUKADA, Masaru (University of Tokyo) [PW, NM]—Prediction of nano-structure by the first-principle electronic state theory
- TSUNEYUKI, Shinji (University of Tokyo—ISSP) [PW]—First-principles study of high-pressure phases of materials
- OGUCHI, Tamio (Hiroshima University) [FLAPW] —First-principles calculation of chalcogens

- OSHIYAMA, Atsushi (Tsukuba University) [PW]—Microscopic mechanisms of epitaxial growth of semiconductors and structural defects in nanometer-scale structures
- SAITO, Susumu (Tokyo Institute of Technology) [PW, etc.]—Physical and material design of hierarchical covalent-network materials
- HAYAKAWA, Kazunobu (Hokkaido University) [FLAPW]—First-principles calculation of exchange interaction between tip and sample surface in exchange force microscope
- FUJIWARA, Takeo (University of Tokyo) [LMTO, LSDA+U, GW]—Electronic structure calculation of condensed matters by the first principle
- HOSHINO, Kozo (Hiroshima University) [PW]—Large-scale molecular dynamics simulation of liquid metals near the critical point
- CHO, Kikuo (Osaka University) [NM]—Microscopic theory of nonlocal response for nanoparticles

Strongly Correlated Electron Systems

This is one of the ISSP's largest research activities, including 60 groups at different Japanese universities and consuming 45% of total computational resources. Investigators in this area focus on effects of strong coulomb repulsions, effects of disorder, electron-phonon interactions in condensed matter. Topics of research and classes of materials along with the number of groups are given below.

- Correlated itinerant electron models (24 groups)
  - Materials in this category include transition metal compounds, high-Tc superconducting cuperates, magnetites, f-electron systems, organic compounds, Hubbard model, t-J model, d-p model, periodic Anderson model, and double exchange model.
- Quantum spin models (19 groups)
  - Physics of Mott insulators, quantum critical phenomena, novel quantum effects and fluctuations, spin excitation gap, ferromagnetic and antiferromagnetic order, and Heisenberg models
- Mesoscopic systems (13 groups)
  - Anderson transition and quantum dots
- Others (5 groups)

Computational techniques for these modeling and simulation efforts use a number of diverse methods. These include exact diagonalization of Hamiltonian matrices using Housholder and Lanczos algorithms; quantum Monte Carlo methods with loop algorithms and auxiliary field for fermions; density matrix renormalization group method at zero and nonzero temperatures; numerical solution of perturbative expansions and self-consistent equations.

Professor Takayama (ISSP) described a new algorithm for problems that are encountered in simulations of spin glasses, polymers, and proteins. These problems are categorized as "slow-relaxation problems in Monte Carlo simulations." These arise due to conflicting constraints and rapid growth of characteristic times as temperature decreases. Efficient algorithms ("exchange Monte Carlo methods") have been developed at SSC-ISSP for addressing this class of difficult but very important problems.

Professor Tsuneyuki (ISSP) reported a very interesting simulation effort using density functional theory for electrons and finite-temperature quantum mechanical simulations for hydrogen. The simulations reported fell in three categories:

1. High-pressure physics—synthesis paths and properties of BC_2N heterodiamond, high-pressure phase of GIC, and atomic SnI_4 at ultra-high pressures
2. Surface science—CO, NO/Pt(111): Blyholder model, halogen-induced etching of silicon surface, and adsorption of organic molecules on Si
3. Quantum distribution of hydrogen atoms in solids—hydrogen and muonium in crystalline silicon, solid hydrogen under megabar pressure

Solid hydrogen has been the "Holy Grail" of electronic structure simulations. Behavior of solid hydrogen is of great interest to astrophysicists and condensed matter physicists. There is indirect evidence of metallization of hydrogen at high temperatures and high pressures and some speculation about geomagnetism of Jupiter. In condensed matter physics, there have been high-pressure experiments using shockwaves and pressure cells. Simulations of solid hydrogen are carried out using DFT (LDA+GGA) to describe electrons and path-integral MD to describe quantum behavior of protons at finite temperatures. This is a major development in simulation techniques of materials where quantum behavior of protons is important. These are rather time-consuming simulations—1 MD step takes 1 teraflop* 1 minute. Distribution of muonium in crystalline silicon at finite temperature has also been studied. Results are in close agreement with experiments.

Professor Imada (ISSP) presented a review of model calculations of highly correlated electronic systems.

Main research interests of the group of *Prof. Saito* (Tokyo Institute of Technology) are in the area of Car-Parrinello-type electronic structure/MD simulations of C_{60}-based condensed-phase systems and doped carbon nanotubes. Results for two important systems were reported:

1. Extensive electronic structure including optimization of atomic positions are carried out for polymerized forms of C_{60}. Questions investigated by Saito's group are why the C_{60} structures polymerize and what the effect of polymerization is on electronic states. They have made predictions based on their simulations of a "new metallic crystalline carbon: the three-dimensionally polymerized C_{60} fullerite."
2. Potassium-doped carbon nanotubes are studied for possible applications as electronic devices. These are medium-size calculations involving 120 atoms per unit cell, where the electronic structure of the pristine nanotube is calculated and atomic positions are optimized and then K is added to the tube, K_2C_{80}—K doped (10,10) rope—and the geometry is optimized again. These simulations are carried out on a single vector-node of the supercomputer. They plan to parallelize their computer code for larger problems.

Professor Watanabe (Science University of Tokyo) described the primary motivation of his project as understanding the microscopic mechanism of field-emitting action of carbon nanotubes using electronic structure simulations and helping design better field emitters. Why carbon nanotubes? A nanotube has high mechanical stiffness and chemical inertness, and it is like an atomic-scale electric wire. The objectives are (1) to clarify the structure-dependent reaction of carbon nanotubes to electric fields, (2) to clarify the role of localized states in field emissions, and (3) ultimately to design ideal nanotube field emitters from first-principles quantum simulations. DFT with LDA+GGA in plane-wave basis sets along with RMM-DIIS (residual minimization method—direct inversion in the iterative subspace) method is used. It is found that electric field is enhanced at the tip and screened inside the thin zigzag-type nanotubes with gapless electronic states. The emission mechanism is governed by the nature of the localized states at the tip.

Professor Tsukada (University of Tokyo) was represented by his former student *Dr. Akagi* at the meeting. (Dr. Akagi is a research associate of Prof. Tsuneyuki.) The purpose of this computational project is to understand in detail the electronic processes involving adsorption, molecular rearrangement, and dissociation of water on silicon (001) surface. On the Si(001) surface, silicon atoms dimerize and form linear chains of dimerized silicons. Experimentally observed saturation coverage of water on Si(001) is less than 0.5 monolayer. This implies that some dangling bonds remain intact on the surface. To investigate the mechanism of dissociation of water in detail, both on-dimer dissociation and inter-dimer dissociation must be permitted. Also, single-molecule and multi-molecular processes should be considered. Energetics for monomer and dimer (of H_2O) on weak and strong sites are evaluated. A detailed adiabatic potential surface of water dimer adsorption system is constructed from the simulation data. The main conclusion of the work is that dissociation of water by multi-molecular process is much more effective.

CONCLUSIONS

The Institute of Solid State Physics is among the best solid-state research institutions in Japan and in the world. The supercomputing facility at the institute provides parallel computing resources to all the university-based researchers in Japan. Approximately 200 groups based at universities use the supercomputer facilities. No commercial codes are used. No industry-based researchers are allowed to

compute at the facility. All the research conducted at the facility is published in prestigious international physics journals. For the first time beginning in year 2000, it may be possible for foreign researchers to use supercomputing facilities at the institute, but only in collaboration with Japanese researchers.

REFERENCES

Akagi, K., and M. Tsukada. 1999. *Surface Science* 438 (1-3) 9.

Hukushima, K., and K. Nemoto. 1996. *J. Phys. Soc. Jpn.* 65 1604.

Imada, M., A. Fujimori and Y. Tokura. 1998. *Rev. Mod. Phys.* 70 1039.

Kitamura, H., S. Tsuneyuki, T. Ogitsu, and T. Miyake. 2000. *Nature* 404 259.

Okada, S., S. Saito, and A. Oshiyama. 1999. *Phys. Rev. Lett.* 83 1986.

Tada, K., and K. Watanabe. 2000. *Jpn. J. Appl. Phys.* 39 268.

Organizations: **Japanese government organizations funding molecular and material modeling:**
Ministry of International Trade and Industry (MITI)
Science and Technology Agency (STA)
Ministry of Education, Science, Sports and Culture (Monbusho)

Author: K. Morokuma

Japanese government grants available in the field of molecular and material modeling are spread through at least three different parts of the government. The following sections describe relevant programs in each organization that involve academic-industry collaboration.

MINISTRY OF INTERNATIONAL TRADE AND INDUSTRY (MITI)

- Program for the Scientific Technology Development for Industries in Collaboration with Academic Institution (*daigaku renkei sangyo kagaku gijutu kenkyu kaihatsu purojekuto*)
  - Budget: $20-25 million/year
  - Length of grants: maximum of 5 years

  Note: This program is run by NEDO (New Energy and Industrial Technology Development Organization, http://www.nedo.go.jp/english/).
- Program for the Scientific Technology Development for Industries That Create New Industries (*shinki sangyou sousei gata teian koubu jigyou*)
  - Purposes: to dig out information buried in academia that may be a seed for industrialization and to support research and development at universities in collaboration with industry
  - Budget: $50-60 million/year
  - One of the projects under this program: "Platform for Designing High Functional Materials" at Nagoya University (the "Doi Project," http://www.stat.cse.nagoya-u.ac.jp/, Principal Investigator: Prof. M. Doi, $4 million/year x 5 years).
  - Pilot program by Prof. K. Hirao (http://www.appchem.t.u-tokyo.ac.jp/labs/hirao/index-e.html) of the Univ. of Tokyo for "Chemical Reaction Simulator" under consideration for a full-scale national project under this program

  Note: This is also run by NEDO.
- Code Development Projects of the Information-Technology Promotion Agency (IPA) http://www.ipa.go.jp/index-j.html
  - IPA, funded by MITI, provides funds for code development. The CAMP project (development of inter-industrial knowledgeware platform by Dr. S. Ohnishi, NEC) and AMOSS project (development of *ab initio* MO software by Dr. T. Takada, NEC) are supported by this program. Those projects are described in the NEC report.

SCIENCE AND TECHNOLOGY AGENCY (STA)

- Fund for Promotion of Science (*kagaku shinkou chousei hi*)
  - Fields: Materials, life sciences, earth-environment
  - Budget: $260 million/year, $1-5 million/grant/year
  - Length, number of grants: 3 + 2 years, 6-8 grants/year

  Note: Several modeling groups are supported by this fund, most of them at MITI institutes but also Prof. A. Miyamoto of Tohoku University.

- Strategic Basic Research Promotion Program (*senryaku teki kiso kagaku suishin jigyou*)
  - Fields: low-environmental-load-type social systems, phenomena under extreme environmental conditions, phenomena in ultra-micro region
  - Budget: $250-300 million/year, $2 million/grant/year
  - Length, number of grants: max. 5 years, 50 grants/year
- Research and Development Applying Advanced Computational Science and Technology (ACT-JST) Program (*koudo keisan kagaku gijutsu kyodou kenkyu suishin jigyou*)
  - Fields: materials, life sciences, environment and safety, earth and astronomical observation
  - Budget: $0.2 million/grant/year
  - Length of Grants: 1 year (short term) to 3 years
  - Web site: http://www-scc.jst.go.jp/cst/ (Japanese only)

  Note: This program is administered by the Japan Science and Technology Corporation (JST). Several molecular and material modeling projects (individual investigators) from academia, national labs and industry are presently supported under this program.

MINISTRY OF EDUCATION, SCIENCE, SPORTS AND CULTURE (MONBUSHO)

- Research for the Future Program (*mirai kaitaku gakujutu kenkyu suishin jigyou*)
  - Fields: science and engineering, life science, complex area
  - Example in Science and Engineering: development of next-generation processes and next-generation materials
  - Budget: $220-250 million/year, $0.5-3 million/grant/year
  - Number of Grants: 110-120 grants/year
  - Note: This program is administered by the JSPS (Japan Society for Promotion of Science). "Computational Science and Engineering" is one of the fields selected as a future research field to be supported under this program. Prof. Yuko Okamoto of the Institute for Molecular Science (IMS) received a grant ($5 million/5 years) entitled, "Development of simulation algorithms for the first-principle prediction of three-dimensional structures of proteins," under this program.

Site:

Joint Research Center for Atom Technology (JRCAT)
Central 4, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8562, Japan
http://www.jrcat.or.jp/

Angstrom Technology Partnership (ATP)
http://www.jrcat.or.jp/rATP/index.html

National Institute for Advanced Interdisciplinary Research (NAIR)
http://www.nair.jp/

Tsukuba Advanced Computing Center (TACC)
1-1-4 Higashi, Tsukuba, Ibaraki 305, JAPAN
http://www.aist.go.jp/TACC/

Date Visited: 12 November 1999

WTEC Attendees: R.S. Jones (report author), S.C. Glotzer, K. Morokuma, E.B. Stechel, P.R. Westmoreland

Hosts:

Dr. Kiyoyuki Terakura, Deputy Project Leader, JRCAT Theory Group; Director, Tsukuba Advanced Computing Center (TACC); Director, Computer Simulation Project; Leader, Computational Science Program; Telephone 81-298-54-2704, terakura@jrcat.or.jp (JRCAT address) or k-terakura@aist.go.jp (after April, '01)
Dr. T. Uda, Project Leader, JRCAT Theory Group
Hiroshi Watanabe, Director for Research Coordination and Planning, NAIR; hwatanabe@nair.go.jp
Dr. Mauro Boero, Chemical Reaction Group, JRCAT
Dr. Ferdi Aryasetiawan, Electronic Excitation Group, JRCAT
Dr. Ikeshouji, TNIRI/NAIR
Dr. Yuto Komeiji, Senior Researcher, Computational Science Group
Dr. Masaaki Kawata, Computational Science Group
Dr. E. Tsuchida, Computational Science Group
Dr. Umpei Nagashima, Computational Science Group, NAIR
Dr. Seiji Tsuzuki, Dept. of Physical Chemistry, National Institute of Materials and Chemical Research (NIMC)

BACKGROUND

The Joint Research Center for Atom Technology (JRCAT) is a research organization located in the buildings of the National Institute for Advanced Interdisciplinary Research (NAIR) in Tsukuba. JRCAT was founded as a joint contract between NAIR and the Angstrom Technology Partnership (ATP) to carry out the Atom Technology Project (which is not referred to as ATP):

- ATP is a consortium of 24 Japanese and 7 foreign companies that was created by the Agency of Industrial Science and Technology (AIST) in February of 1993.
- NAIR was also created by AIST in January of 1993 and operates research projects of 5-10 years in duration.
- The Atom Technology Project is one of six current NAIR projects. It was begun in January 1993 with a total budget of about ¥25 billion as a two-phase project for a planned period of 10 years. Funding is 90% from NEDO (through ATP) and 10% from NAIR. Phase II began in April 1998 with some modification of staff members and group structures. The project is open to research scientists from industrial, academic, and government sectors with the expectation that the group would be disbanded and researchers return to their home institutes when the project is finished. JRCAT's project life is supposed to be completed on March 31, 2002.

The Theory Group headed by Drs. Terakura and Uda is one of approximately 10 research groups currently at JRCAT. This group consists of 18 research scientists and 5 engineers to support the JRCAT supercomputer system. Among the research scientists, ten are post-docs, only 3 of whom are Japanese. The activities are focused about 80% on *ab initio* molecular dynamics of the Car-Parrinello type. (This is separate from the multi-industry CAMP activity described at NEC.) JRCAT's supercomputer system was upgraded in February 1999 to a Fujitsu VPP700E/20 with 21GB main memory and 180GB disk storage and an IBM RS/6000 SP/48 with 24GB main memory and 285GB disk storage.

Dr. Terakura is also group leader of a new NAIR project, the Computer Simulation (CS) Project, begun in April 1999. This is a 6-year project, devoted solely to computer modeling. Stated goals are to develop new theories and algorithms and apply them to predict complicated static and dynamic phenomena in areas including engineering, biotechnology, macro structures and global environmental issues. Current group members are K. Terakura (specializing in condensed matter physics), U. Nagashima (quantum chemistry, formerly at IMS), Y. Komeiji (biological applications including MD of proteins), Y. Morikawa (first-principles chemistry using plane-wave MD), M. Kawata (algorithms for Monte Carlo and MD), and E. Tsuchida (materials modeling). Application research will be performed primarily by members of the Computational Science Program, a group of approximately 50 scientists from 14 AIST institutes. They will use a network linking them to the central research.

In a third role, Dr. Terakura is Director of the Tsukuba Advanced Computing Center (TACC). The center is run by AIST and includes approximately one Tflop of computing power, available to all of the AIST laboratories:

- A Hitachi SR8000/64 with 512GB main memory and a peak performance of 512 Gflops
- An IBM RS/6000/128 with 256GB main memory and peak performance of 200 Gflops
- A custom-made Compaq Alpha cluster with 256 nodes with 256GB main memory and a peak performance of 256 Gflops

In April of 2001, AIST planned to restructure its 15 institutes into a single organization consisting of 22 research institutes and 23 research centers. Dr. Terakura described a "research institute" as a more-or-less permanent organization representing a fundamental and generic area. A "research center" will be set up similar to the NAIR project, with a finite lifetime of approximately 5-10 years and a concentrated investment of resources. Some members of the JRCAT theory group and the whole CS group were scheduled to join the Research Institute for Computational Sciences (RICS).

R&D ACTIVITIES

WTEC panelists heard presentations by two members of JRCAT, four members of the CS group, and by one member of the National Institute of Materials and Chemical Research, another AIST laboratory located at the JRCAT site.

The theory group at JRCAT divides its activities into four categories: Surface Processes, including AFM, STM and nanostructures; Chemical Reactions, focusing on catalysis on surfaces and in solution; Phase Control, investigating transition-metal oxides and organic solids; and Electronic Excitation, focusing on spectroscopy.

Dr. Boero of the Chemical Reactions group presented results from his work on Ziegler-Natta heterogeneous catalysis of ethylene to polyethylene (Boero, Parrinello, and Terakura 1998; Boero, Parrinello, and Terakura 1999). His background includes a doctorate with Roberto Car at Lausanne and a postdoc in Michele Parrinello's group. He has worked out the energetics and reaction pathway using the Car-Parrinello molecular dynamics approach with a generalized gradient approximation to density functional theory. In particular, C_2H_4 adsorption onto $TiCl_4$ on $MgCl_2$ was examined for two different Ti configurations.

Dr. Aryasetiawan of the Electronic Excitation group presented work on first-principles calculations of excited state properties in solids. He is developing methods to extend the GW approximation and showed applications to YH_2-YH_3 systems that transform from a metal to an insulator as H is absorbed (Miyake,

Aryasetiawan, and Terakura 2000). He also discussed a Green's function formalism for calculating spin-wave excitations and gave results for Ni and Fe (Karlsson and Aryasetiawan 2000).

Members of the CS group have multiple affiliations, presumably because the group is so new. Dr. Ikeshouji is also a member of the Tohoku National Industrial Research Institute (TNIRI), another AIST institute. He is developing a general purpose molecular dynamics code for large scale systems that will include databases of molecular structures and pseudopotentials, quantum and classical MD codes, biological codes (PEACH) and visualization and analysis tools. It will use the TACC for its main computing engine, but all tools will be available through a sophisticated GUI from Web terminals. The expected completion date is March 2002.

Dr. Komeiji is also associated with two other AIST institutes: the Electrotechnical Laboratory (ETL) and the National Institute of Bioscience and Human Technology (NIBH). He presented a summary of biomolecular simulation work being done in various AIST institutes. This is reproduced in Table B2.4 below. He described his own work developing a classical MD biological system simulator, called PEACH for "Program for Energetic Analysis of bioCHemical molecules." This program (available through the Web page http://www.nair.go.jp/comp/soft.htm) will be included in the code package described by Ikeshouji.

Table B2.4
Biomolecular Simulation and Computing in AIST

Approach	Scientist	Location	Software	Problems
Classical MD	Y. Komeiji M. Uebayasi	NAIR/ETL/NIBH NIBH	PEACH (in house)	Protein dynamics DNA/Peptide interactions Software development
	M. Tateno	NIBH	AMBER	DNA/Peptide interactions
	M. Iwakura	NIBH	AMBER	Molecular modeling of mutant proteins
Quantum MD	M. Boero	JRCAT	CPMD	Ribozyme
Molecular orbital calculation	M. Uebayasi T. Uchimaru K. Taira	NIBH NIMC NAIR	GAUSSIAN	Ribozyme
Genome informatics	K. Asai	ETL		Prediction of protein 3D structure
Molecular graphics	Y. Ueno	ETL	MOSBY (in house)	Graphics software Electron Microscope image construction

Source: Yuto Komeiji

Dr. Kawata presented ideas to speed up classical MD calculations. These include parallelization of the particle mesh Ewald and the multipole methods. Benchmark calculations suggest he may be able to speed up his calculations by 3 orders of magnitude.

Dr. Tsuchida discussed using the finite element method for density functional calculations to obtain an order(N) method. Because the method lends itself to efficient parallelization, he hopes to obtain significant speed improvements.

Dr. Tsuzuki from NIMC gave the final presentation. He is performing classical MD calculations on small molecules using *ab-initio*-based force fields. He described a number of different calculations on various systems, all of which have been published recently (Tsuzuki, Uchimaru, and Tanabe 1998; Tsuzuki et al. 1996; Tsuzuki et al. 1998; Tsuzuki, Uchimaru, and Tanabe 1998; Tsuzuki, Klopper, and Lüthi 1999; Tsuzuki et al. 1999). Careful comparisons with data are especially important.

Dr. Nagashima from the Computer Simulation Project was also at the meeting briefly near the end and did not give a presentation. He contributed a reprint of an article describing his "Molecular Orbital Calculation Engine" (MOE) (Nagashima et al. 1999), which is a special-purpose computer for MO calculations. The chosen software is GAMESS-US.

The collaboration of Prof. Terakura with Ivan Stich (Slovak Technical University, http://kf-lin.elf.stuba.sk/ccms/) on modeling atomic-force microscopy should be noted (Ke at al. 1999; Tóbik et al. 1999).

DISCUSSION

Collaborations are very important at JRCAT, both with Japanese companies and internationally. It helps that if a researcher from an ATP member company comes to work at JRCAT, funding is 70% from the government. Despite the materials focus, Dr. Nagashima noted that chemical companies had been especially responsive. He speculated that applicability is better or more natural for the chemical companies, who have long experience in the field.

REFERENCES

Boero, M., M. Parrinello and K. Terakura. 1999. Ziegler-Natta heterogeneous catalysis by first principles computer experiments. *Surf. Sci.* 438, 1-8.

Boero, M., M. Parrinello, and K. Terakura. 1998. First principles molecular dynamics study of Ziegler-Natta heterogeneous catalysis. *J. Am. Chem. Soc.* 120, 2746-2752.

Karlsson, K., and F. Aryasetiawan. 2000. Spin-wave excitation spectra of Nickel and Iron. *Phys. Rev. B* 62, 3006-3009.

Ke, S.H., T. Uda, R. Pérez, I. Stich, and K. Terakura. 1999. First-principles investigation of tip-surface interaction on a GaAs(110) surface: Implications for atomic force and scanning tunneling microscopies. *Phys. Rev. B* 60, 11631-11638.

Miyake, T., F. Aryasetiawan, and K. Terakura. 2000. GW quasiparticle band structure of YH_3. *Phys. Rev. B* 61, 16491-16496.

Nagashima, U., S. Obara, K. Murakami, T. Yoshii, S. Shirakawa, T. Amisake, K. Kitamura, O. Kitao, and K. Tanabe. 1999. Development of MOE (molecular orbital calculation engine). *Computational Material Science* 14, 132-134.

Tóbik, J., I. Stich, R. Pérez, and K. Terakura. 1999. Simulation of tip-surface interactions in atomic force microscopy of an InP(110) surface with a Si tip. *Phys. Rev. B* 60, 11639-11644.

Tsuzuki, S., T. Uchimaru, and K. Tanabe. 1998. Ab initio calculations of intermolecular interaction potentials of corannulene dimer. *J. Phys. Chem. A* 102(4), 740-743.

Tsuzuki, S., T. Uchimaru, and K. Tanabe. 1998. Intermolecular interaction potentials of methane and ethylene dimers calculated with the Møller-Plesset, coupled cluster and density functional methods. *Chem. Phys. Letters* 287(1/2), 202-208.

Tsuzuki, S., T. Uchimaru, K. Matsumura, M. Mikami, and K. Tanabe. 1999. Effects of basis set and electron correlation on the calculated interaction energies of hydrogen bonding complexes: MP2/cc-pV5Z calculations of H_2O—MeOH, H_2O—Me_2O, H_2O—H_2CO, MeOH—MeOH, and HCOOH—HCOOH complexes. *J. Chem. Phys.* 110, 11906-10.

Tsuzuki, S., T. Uchimaru, K. Tanabe, S. Kuwajima, N. Tajima, and T. Hirano. 1996. Refinement of nonbonding interaction parameters for carbon dioxide on the basis of the pair potentials obtained by MP2/6-311+G(2df)-level ab initio molecular orbital calculations. *J. Phys. Chem.* 100(11), 4400-4407.

Tsuzuki, S., T. Uchimaru, M. Mikami, and K. Tanabe. 1998. New medium-size basis sets to evaluate the dispersion interaction of hydrocarbon molecules. *J. Phys. Chem. A* 102(12), 2091-2094.

Tsuzuki, S., W. Klopper, and H.P. Lüthi. 1999. High-level ab initio computations of structures and relative energies of two isomers of the CO_2 trimer. *J. Chem. Phys.* 111, 3846-54.

Site: **Mitsubishi Chemicals, Yokohama Research Center**
Mitsubishi-Tokyo Pharmaceuticals Inc.
1000, Komashida, Aoba-ku
Yokohama 227, JAPAN
Telephone: (81) 45-963-3011; Fax: (81) 45-963-3970
http://www.m-kagaku.co.jp/index_en.htm

Date Visited: 8 November 1999

WTEC Attendees: P.T. Cummings (report author), S.C. Glotzer, R. Miranda, K. Morokuma, E.B. Stechel, P.R. Westmoreland

Hosts: Yukikazu Natori, General Manager, Research Planning Department, 2206791@cc.m-kagaku.co.jp
Dr. Takeshi Ishikawa, General Manager, Research and Development Coordination Department, 3206014@cc.m-kagaku.co.jp
Dr. Hiroaki Ueno, Senior Research Scientist, Pharmaceuticals Discovery Laboratory, 1303612@mitsubishi-pharm.co.jp
Dr. Hideo Kubodera, Mitsubishi-Tokyo Pharmaceuticals Inc., 1607753@mitsubishi-pharm.co.jp
Dr. Shinichiro Nakamura, Senior Research Scientist, Computer Science Laboratory, shin@rc.m-kagaku.co.jp
Dr. Hisao Takeuchi, Senior Research Scientist, Computer Science Laboratory, take@rc.m-kagaku.co.jp

BACKGROUND

Mitsubishi Chemicals is a diverse chemicals company producing petrochemicals, carbon products, agrochemicals, information- and electronics-related products, pharmaceuticals, plastics-based products and specialty chemicals. Annual sales in Japan are approximately $12.8 billion/year, of which petrochemicals account for $3.6 billion/year and carbon and agrochemicals $1.8 billion/year. Included in the remaining $7.4 billion/year of sales is $0.8 billion/year in pharmaceuticals sales. Mitsubishi Chemicals generates an additional $2 billion/year in sales outside Japan. The company makes over 20,000 products.

R&D ACTIVITIES

Mitsubishi Chemicals

Research and development activity at Mitsubishi Chemicals is conducted at five research facilities located at Yokkaichi, Yokohama, Kurosaki, Mizushima and Tsukuba, resulting in 1600-1800 patents per year. The Yokohama Research Center is the corporate central research laboratory of Mitsubishi Chemicals. It employs 1,500 people, of whom 800 are researchers. The annual research budget of the Yokohama Research Center is around $200 million/year. About $10 million/year is expended on outside research at universities. Approximately 800 patents per year originate from the Yokohama Research Center.

The size of the modeling and computational chemistry effort at Mitsubishi Chemicals is substantial, with five researchers in the pharmaceuticals area and eleven in the non-pharmaceuticals area. Several of the non-pharmaceuticals modelers in Yokohama are part of the Computer Science Laboratory, a group of 20 people with a budget of about $5 million.

Molecular modeling and computational chemistry activities at Mitsubishi Chemicals support the following:

- Design and prediction of properties of chemicals and materials (e.g., catalysts, dyes and polymers)
- Prediction of process properties for design and optimization
- Design and optimization of pharmaceuticals and the development of processes to manufacture them

The tools used are diverse: wavefunction-based methods, electronic density functional theory, semi-empirical molecular orbital theory, molecular mechanics, atomistic/molecule/mesoscale simulation, hybrid quantum mechanical/molecular mechanics methods, QSAR, QSPR, chemoinformatics, bioinformatics and data-mining. About 90% of the calculations are performed with commercial software, the remaining 10% being in-house developed. The software packages used are: MSI, Gaussian, MOPAC, parallel GAMESS-UK, Wien97, Car-Parrinello molecular dynamics (CPMD from IBM-Zurich), BOSS, CAMOD, DOCK, COMFA (a QSAR package). Gaussian is the most frequently used package. For mesoscale modeling, Mitsubishi's modelers use an in-house program based on the lattice self-consistent field theory.

The hardware used to perform calculations includes SGI and DEC workstations, PC clusters running FreeBSD, Fujitsu VX (a small vector computer), Cray C90, SGI Origin 2000, and IBM SP/2.

The researchers at Mitsubishi Chemicals estimated that around 5% of the patents originating from the Yokohama Research Center make use of molecular modeling and computational chemistry methods. One computational chemist, Dr. Nakamura, holds 20 to 30 patents based on his involvement in product development. Pharmaceuticals patents involving computational chemists are expected to increase as a consequence of a new policy to include a modeler in every pharmaceutical project. Publication of non-proprietary research results in the peer-reviewed literature is encouraged.

In the general chemical area, the applications for molecular modeling have included the molecular design of transition-metal catalysts, solid-state catalysts, dyes and pigments, photonic materials, battery-related materials, colloidal dispersions, and additives; molecular modeling has been used to assist in solvent selection and to predict physical properties.

Successes. A particular success has been the use of molecular modeling tools to design homogeneous metallocene catalysts. Now, the catalysis people are doing most of their own calculations.

A second example is the development of new polymer membranes. In 1988-1990, researchers at Mitsubishi Chemical performed molecular simulations of penetrant diffusion in polymers using a molecular dynamics code developed in-house. The understanding that followed from these simulations led to the development of a new polymer membrane. This was the first molecular simulation success at Mitsubishi Chemical.

Still another example is the development of new dyes, described by Dr. Nakamura. Dyes are key compounds in many high-technology devices and products such as CD-R, DVD, and xerography. An experimental team at Mitsubishi Chemicals established by X-ray the structure of one near-IR-absorbing dye, showing that the dye molecule had a twisted structure (Kubo et al. 1989). This was unexpected. According to conventional wisdom, near-IR absorbing dyes must be mostly planar, since near-IR absorption is associated with an extended—i.e., planar—configuration. Using computational chemistry methods, the theoreticians showed that the configurations giving rise to the near-IR absorption were $\bullet\text{-}\bullet^*$ and $n\text{-}\bullet$ mixing. This yielded insight to the experimentalists, who were then able to search a wide range of structures. Some of this work has been reported in the open literature (Adachi, Murata, and Nakamura 1993). While performing this research, the theoreticians at Mitsubishi Chemicals needed to calibrate for themselves the methodologies used (ZINDO, which later was commercialized by CAChe). At the time this work was being done, there were no tested data. The need to use ZINDO methods was a result of the dye molecules being too large for their structures to be calculated by *ab initio* excited-state molecular-orbital methods (Adachi and Nakamura 1991). Current work applies time-dependent DFT, making use of Gaussian.

Fourth, a surfactant problem was the first effective application of mean-field methods for polymers at Mitsubishi. Combinations of mono- and di-esters were to be used. Experimentally, critical micelle concentrations were measured for selected model compounds. Micelle-micelle interactions were modeled for

the mixed surfactants system, and the new understanding resulted in a QSPR relation that proved to be a valuable guide.

A fifth application was development of dispersants. Particles were to be coated with polymers having a complicated architecture to prevent the particles from adhering to one another. Simulation was used to identify the optimal polymer architecture. An unexpected bridging effect was identified that aided qualitative understanding of the phenomenon as well.

Mitsubishi-Tokyo Pharmaceuticals Inc.

The visit to Mitsubishi Chemicals was unusual because Mitsubishi also has a large pharmaceuticals effort. There are five modelers working on pharmaceutical problems. There is one modeler involved in every pharmaceutical project. In general there are one to two projects per modeler.

A key tool is in-house software called MCC CADD (Computer-Aided Drug Design). This combined combinatorial and computational approach is very effective. One exemplary project was a search for an enzyme inhibitor. Initially, a 1-D protein sequence and a 3-D structure database were used as a basis for homology modeling. These results were used in turn to identify the active site. Using this information, a 66,000-compound-ligand database, and the compound-screening program DOCK, 44 compounds were selected as having a high probability of success, of which eighteen passed assay tests for activity. Further tests for cell toxicity yielded two suitable compounds. The 41% yield of successful activity hits is far superior to the 2% identified by a typical, purely random combinatorial experimental screening.

Relationship to other "technologies" (X-ray, NMR, bioinformatics, chemical informatics, genomics) in the company. There is an X-ray crystallographer in the pharmaceutical division and an X-ray group for X-ray and NMR that is company-wide. Mitsubishi has joined the JBIC (Japan Bioindustry Consortium), which is working on structural and functional genomics, and the company plans to utilize that information in its research. One person has responsibility for bioinformatics and chemical informatics as part of the modeling group in pharma. They use informatics for QSAR, analysis of molecular diversity, and prediction of absorption of compounds. They do have an effort to predict sites of metabolism and adsorption and are trying to build models to predict absorption and excretion. There is an active combinatorial chemistry research program in the company, which the modelers participate in, in order to help select appropriate building blocks for the molecules, based on their physicochemical properties, and to evaluate diversity of the compounds.

Relationship to chemists. The modeling in pharmaceuticals seems to be appreciated by the medicinal chemists, as witnessed by the involvement of a modeler in each project.

Credit and how success is evaluated. There is some inclusion of modelers in patents in pharmaceuticals and this involvement will likely increase as there is more involvement of modelers in each project.

Software/hardware capabilities and needs. The main software used in pharmaceuticals modeling are DOCK and COMFA(3D-QSAR).

GENERAL DISCUSSION

Mitsubishi Chemicals participates in several national government projects. One, sponsored by the Japan Science and Technology Agency (JSTA), is aimed at developing modeling tools for heterogeneous and homogeneous catalysis. Another JSTA project is on photochromic materials. Another is the NEDO Doi Project, described in detail elsewhere in this report. Other projects include the Japan Biotechnology Consortium and the NEDO Bioinformatics Consortium.

Mitsubishi was also a charter member of the MSI (Biosym) Polymerization Consortium. It presently also participates in the catalysis/adsorption and materials science consortia, finding this approach valuable both for the codes and for sharing of information.

CONCLUSIONS

The molecular modeling efforts at Mitsubishi Chemicals are mature, well integrated into the broader research and development activities, and rewarded and recognized by the company (e.g., through patent recognition). Paradoxically, the researchers' perception was that these are not yet recognized as powerful tools throughout the entire company. They were nevertheless confident that acceptance is becoming wider. As Dr. Takeuchi put it, probably the biggest key is continuing to examine how to solve problems, not just working to develop the problem solving tools.

REFERENCES

Adachi, M., and S. Nakamura. 1991. *Dye and Pigment* 17, 287.

Adachi, M., Murata, and S. Nakamura. 1993. *JACS* 115, 4331.

Kubo, Y., et al. 1989. *Chem. Comm.* 35.

Site: **NEC Tsukuba Laboratories**
34, Miyukigaoka
Tsukuba, Ibaraki 305-8501, Japan
http://www.nec.com

Date Visited: 11 November 1999

WTEC Attendees: P.R. Westmoreland (report author), A.M. Chaka, P.T. Cummings, S.C. Glotzer, R.S. Jones, P.A. Kollman, R. Miranda, K. Morokuma, H. Morishita, E.B. Stechel, P. Vashishta

Hosts: Dr. Hitoshi Igarashi, Chief Engineer, R&D Group; igara-h@frl.cl.nec.co.jp [In 2000: Chief Engineer, System Devices and Fundamental Research; 4-1-1 Miyazaki, Miyamae-ku, Kawasaki, 216-8555 Japan; Phone (81) 44-856-8164, FAX (81) 44-856-8165, igara-h@rdg.cl.nec.co.jp]
Dr. Akihiro Iwaya, Executive Expert, High-Performance Computing, Supercomputers Marketing Promotion Division; iwaya@sxsmd.ho.nec.co.jp
Dr. Toshikazu Takada, Research Fellow, Fundamental Research Laboratories; takada@frl.cl.nec.co.jp
Dr. Yoshiyuki Miyamoto, Principal Researcher, Fundamental Research Laboratories
Dr. Shuhei Ohnishi, Research Fellow, Fundamental Research Laboratories; ohnishi@frl.cl.nec.co.jp
Akinori Yamamoto, Assistant Manager, NEC Informatec Systems, Ltd., Shared Infrastructure and Services Division; yamamoto@pb.jp.nec.com

BACKGROUND

NEC is one of the world's largest companies in the areas of information and communications systems, having 1998 sales of ¥4.8 trillion ($40 billion). Products include communications systems and equipment, computers from PCs to supercomputers, software, lasers, semiconductors, electronic components, and televisions. It employs 158,000 people around the world and operates 125 manufacturing plants in 76 countries.

To keep its products at the cutting edge, NEC operates 13 R&D facilities, seven of which are in Japan, three in the United States, and three in Germany. Dr. Igarashi described the Tsukuba Research Laboratories as NEC's center for advanced basic research. Its research includes the development of molecular modeling for solving NEC's application problems and for use on NEC computers by its customers. The Tsukuba Laboratories employ 230 researchers, about 20 of whom are modelers of different kinds.

R&D ACTIVITIES

High-Performance Computing

Dr. Iwaya discussed current and future NEC high-performance computing, expressing its effective use as a key goal for the company. Its market was described as computers at a medium range of affordability ($100 thousand-$10 million) and high end (>$10 million), as opposed to "commodity" PCs and workstations. Their vision is that both these higher sectors would use UNIX on scalable parallel vector microprocessors rather than purely scalar microprocessors operated in parallel configurations. The argument was based on memory bandwidth and memory latency being limiting for efficiency, with the vector architectures having much higher bandwidth and latency. Further differentiation in cost and power would come from the use of shared vs. distributed memory.

In terms of current top-end performance, NEC's SX-5 supercomputer can perform to 3 teraflops, maintaining 25 to 40% efficiency on up to 40 parallel vector processors. (This is a shared-memory-processor machine, with 15 delivered at the end of the first half of FY99.) Compared to the 1975 Cray-1, this system is roughly

30,000 times faster. In the near-term, they project their high-performance supercomputers to use either shared-memory clusters of nodes (each of which would be a multi-processor, shared-memory machine) or a (cc)-NUMA architecture (cache-coherent, non-uniform memory access). The latter would use large caches of dedicated memory for each processor, coordinated so all the memory appears shared.

The largest project discussed was a whole-earth weather simulator, which is scheduled to be completed in the first quarter of 2002 for the Japan Science & Technology program. The "Earth Simulator" would be 640 SX-4 nodes, delivering 40 teraflops peak performance. When taken with its ancillary components, it will fill two floors of a 65 m x 50 m building.

Developing Quantum Chemistry

Dr. Takada described the dual activities of the quantum chemistry group, developing new codes for research purposes and dedicated special-purpose computers.

The principal code-development activity was AMOSS (*ab initio* Molecular Orbital System for Supercomputers), aimed at performing quantum chemistry for protein conformation and activity. Its key is an algorithm for parallel generation of the integrals in the Fock matrix. A calculation on kinase C was the largest shown, involving 856 atoms and 8672 basis functions (the MIDI-4 + polarization basis set). Giving more specific performance figures, calculating energy of a $C_{72}H_{72}N_8O_{12}Mg_2$ complex involving 986 atomic orbitals required on the order of one computer hour operating at 40 gigaflop (SX-4/20 computer).

In a practical environment, such calculations require a balanced view of supercomputers and networks involving PCs and workstations. Web interfaces on an intranet were proposed as a solution. A new NEC-developed code for Web-based submission and viewing, VRMS (virtual microscope), was demonstrated. The two codes are part of a national project for making software in Japan, now complete.

An alternative approach is to produce very fast special-purpose chips. A 560 Mflop prototype for evaluating s-orbital Fock integrals has been developed with Altera Corporation of San Jose (http://www.altera.com/).

Developing Applications

Dr. Miyamoto described several applications that have been examined by the Fundamental Research Laboratories, summarized in Table B2.5. These applications focus on a variety of problems ranging from electronic properties of semiconductors (Si, GaAs, GaN) to insulators (SiO_2) to environmental concerns (dioxins, perfluorocarbons). Molecular-orbital theory and electronic density-functional theories at local-density approximation, GGA, and hybrid levels are used to develop physical and chemical properties, sometimes being used in process simulations.

Problems are usually selected because they promise results that will help solve present or future problems, help in devising experiments, or develop methods or results that showcase the power of NEC computers.

There is some collaboration with experimental researchers, such as nanotube research and materials growth processes.

Although the Fundamental Research Laboratory is set up on an academic or "Bell Labs" model, its researchers also cooperate with process and manufacturing parts of the company. An example is the problem of dopant diffusion in silicon, a common issue for manufacturing. Temperatures must be low enough to minimize diffusion, or incident particle energy for MBE must be low enough. Process personnel can perform macroscale simulations, but they requested FRL to calculate certain diffusion barriers. Because NEC's operations are spread out, such requests usually come in by phone or e-mail. When such work is done with internal clients, there is usually some working together to decide the best objectives and approach.

Table B2.5
Technical Summary of Molecular and Materials Modeling at NEC Tsukuba, FRL

Past and Present Applications
Reconstruction of CoNi interface between tungsten interconnect and silicon, compared to images from transmission electron microscopy (Yu et al. 1999)
Diffusional penetration of dopants (Yu et al. 1998)
$B_xC_yN_z$ nanotubes, including nanowires with chiral currents (Miyamoto et al. 1994; Miyamoto et al. 1996a; Miyamoto et al. 1996b)
Melting of silicon, carbon, and $TiSi_2$ by *ab initio* molecular dynamics (Sugino and Car 1995; Sugino 1999)
Reactions on excited potential energy surfaces to examine H deactivation of Si donor in GaAs, H on Si (111), H trapped at vacancies in SiO_2 (Sugino and Miyamoto 1999a; Miyamoto, Sugino, and Mochizuki 1999a; Miyamoto, Sugino, and Mochizuki 1999b)
Catalytic destruction of dioxin by dissociative adsorption on silicon, breaking -O- linkages (Okamoto 1999a; Okamoto 1999b)
Predictions of electrical conductivity (Sugino and Miyamoto 1999b)

Participation in the CAMP Project

Dr. Ohnishi presented the multi-industry CAMP project, Collaborative Activities for Material Science Programs (http://www.camp.or.jp), which he leads. Representatives from ten companies (Table II.6) cooperate in research on materials-modeling subjects in which they have broadly overlapping interests:

- First-principles molecular dynamics: developed "CAMP-Atami" code along the lines of Car-Parrinello in 1990-1994. It is freely distributed by JCPE (Japan Chemistry Program Exchange, http://www.aist.go.jp/NIMC/TC/~jcpe). Current work seeks to exploit the databases and feed into the simulation needs described below.
- Database of Free Atoms: pseudopotentials, initial charge distribution, and wavefunctions incorporating relativistic effects. Pseudopotentials have been developed up to atomic number 103 (Lr).
- Database for "Canonical atoms": from simulations seeking self-consistent electron distribution of molecules, clusters, and solids for any atomic arrangement/composition, obtain distribution functions $\rho_{SCF}(r) = \Sigma [\rho_{CANONICAL\ ATOM}(r - R_{CA})]$; examples to date include Au fcc structure and SiO_2 alpha-quartz structure.
- Advanced simulation systems: mixed basis sets (plane-wave and localized atom models) and system-adapted generalized coordinates.
- High-performance computing: includes developments for parallel supercomputers, massively parallel and distributed.
- Development of knowledge-ware: user interfaces and data structures for information exchange. The present focus is to use XML structures within Java, set up within a Web-browser type of interface (Figure B2.1).

The program is now in its third phase. The first phase was purely voluntary and resulted in the CAMP-Atami code. In the next phase (1994-98), additional software was developed and distributed, supported over two years with 500 million yen from the Information-technology Promotion Agency of Japan (IPA/MITI). The current phase (1998-2000) reaches more broadly on the "knowledge-ware" project, and is supported by IPA funding (¥200 million). Again, the codes will be distributed as free software. The CAMM activity (Table B2.6) does not develop codes with government support, but is a potential group of users who are kept aware of the code developments.

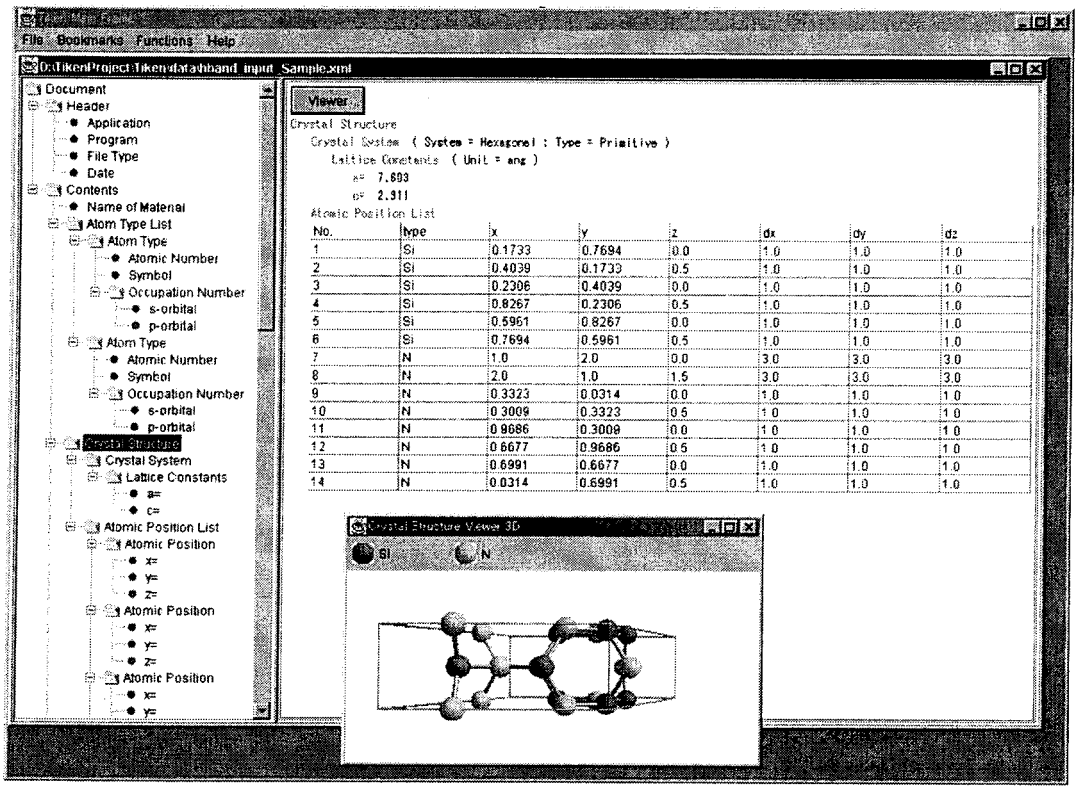


Fig. B2.1. Example of XML tags and browser interface.

Table B2.6
Industrial Participation in CAMP and CAMM Cooperative Programs

CAMP Project (partially supported by IPA/MITI)	CAMM Forum (supported by Business Research Institute)
NEC Corp., Fundamental Research Lab.	NEC Informatic Systems, Ltd.
Toyota Central R&D Labs Inc.	Canon Inc.
Sumitomo Chemical Co. Ltd., Tsukuba Research Lab	Seiko Epson Corp.
Asahi Glass Co. Ltd., Research Center	Nippon Steel Corp.
Sony Corp., Research Center	TDK Corp.
Toshiba Corp., Research and Development Center	Murata Manufacturing Co., Ltd.
Pioneer Electronic Corp., Corporate Research and Development Lab	Central Research Institute of Electric Power Industry
FDK Corp., Research and Development Div.	Ishikawajima-Harima Heavy Industries Co., Ltd.
Mitsubishi Heavy Industries, Ltd., Advanced Technology Research Center	National Institute of Materials and Chemical Research
Fuji Research Institute Corp.	

CONCLUSIONS

NEC is involved in a range of basic and applications research, focused substantially on materials. Its employees mainly use *ab-initio*-based methods, either electronic DFT or *ab initio* molecular dynamics, on parallel platforms. However, there is also significant method and parameter development, both internally and as a leader in the CAMP multi-industry program.

REFERENCES

Miyamoto, Y., A. Rubio, S.G. Louie, and M.L. Cohen. 1994. *Phys. Rev.* B 50, 4976.

Miyamoto, Y., A. Rubio, S.G. Louie, and M.L. Cohen. 1996a. *Phys. Rev. Lett.* 76, 2121.

Miyamoto, Y., A. Rubio, S.G. Louie, and M.L. Cohen. 1996b. *Phys. Rev. B* 54, R11149.

Miyamoto, Y., O. Sugino, and Y. Mochizuki. 1999a. First-principle dynamics of defect reactions triggered by electronic excitation. *Physica B* 273-274, 991-994.

Miyamoto, Y., O. Sugino, and Y. Mochizuki. 1999b. Real-time electron-ion dynamics for photoinduced reactivation of hydrogen-passivated donors in GaAs. *Appl. Phys. Lett.*, 75, 2915-2917.

Okamoto, Y. 1999a. *J. Phys. Chem. A*, 103, 7686-7691

Okamoto, Y. 1999b. *Chem. Phys. Letters*, 310, 395.

Sugino, O. 1999. First-principles calculations of the phase diagram of Si. Paper C25.01, American Physical Society, San Francisco CA (20 Mar.); http://flux.aps.org/meetings/BAPSMAR95/abs/SC2501.html.

Sugino, O., and R. Car. 1995. *Phys. Rev. Letters*, 94, 1823.

Sugino, O., and Y. Miyamoto. 1999a. Density-functional approach to electron dynamics: Stable simulation under a self-consistent field. *Phys. Rev. B*, 2579-2586 (15 Jan.).

Sugino, O., and Y. Miyamoto. 1999b. Real-time density functional approach to the electric conductivity. Paper EC10.05, American Physical Society, Atlanta GA (22 Mar 1999); http://www.aps.org/meet/CENT99/BAPS/abs/S1565005.html.

Yu, B.D., et al. 1998. *Phys. Rev.*

Yu, B.D., Y. Miyamoto, O. Sugino, T. Sasaki, and T. Ohno. 1999. Materials Research Society, San Francisco CA (April).

Site: **Nagoya University**
Ohmine Research Group
Department of Chemistry, Faculty of Science
Furo-cho, Chikusa-ku
Nagoya 464-8602, Japan
Telephone: +81-52-783-7036; Fax: +81-52-783-7036
URL: http://www.chem.nagoya-u.ac.jp/ohmine/prf.html

Date Visited: 9 November 1999

WTEC Attendees: P.T. Cummings (report author), S.C. Glotzer, R. Miranda, K. Morokuma, E.B. Stechel, P.R. Westmoreland

Hosts:

Dr. Iwao Ohmine, Professor, Department of Chemistry, ohmine@aqua.chem.nagoya-u.ac.jp
Dr. Shinji Saito, Associate Professor, Department of Chemistry, shinji@chem.nagoya-u.ac.jp

BACKGROUND

The research group headed by Professor Iwao Ohmine is one of the best-known simulation groups in Japanese academia. It currently consists of 17 group members, most of whom are graduate and undergraduate students at the university, an associate professor (Dr. Saito), and a full professor (Dr. Ohmine). The group size and structure are typical of other groups at Nagoya University Chemistry Department, which is one of the best-funded chemistry departments in Japan.

R&D ACTIVITIES

The group's research focus is on water, aqueous solutions, and biomolecules. The research is very fundamental, with no industrial interaction or focus. All software codes (classical MD simulation codes, analysis codes, visualization codes) are developed in the group or based on public-domain software (e.g., the public domain POVRAY rendering software). The research is supported by substantial computational resources, including approximately 50 DEC alpha or SGI workstations operating as stand alone machines. The group recently received a $2.5 million equipment grant that enabled them to expand their computational resources substantially. They do not use parallel computers.

A problem of central interest in the Ohmine research group is the dynamics of liquid and low-temperature water, particularly the hydrogen bond dynamics. A review of their work has recently published (Ohmine and Saito 1999).

Because it is the most important solvent in nature (e.g., the solvent in which all biochemistry takes place), water continues to be a subject of intense research. Water remains perhaps the least understood of all low-molecular-weight fluids (because of a subtle balance between the various interactions between water molecules giving rise to complex and atypical properties). The Ohmine research group is one of many throughout the world that focus all or a substantial part of their efforts on trying to understand the molecular basis of water's many unusual properties.

CONCLUSIONS

In summary, the Ohmine group at Nagoya University is an example of an excellent academic research group focused on fundamental science. It is typical of scores of other such research groups around the world using and advancing the state of the art of molecular and materials modeling.

REFERENCE

Ohmine, I., and S. Saito. 1999 Water dynamics: fluctuation, relaxation, and chemical reactions in hydrogen bond network rearrangement. *Acc. Chem. Res.* 32. 741-749.

Site: **RIKEN—The Institute of Physical and Chemical Research**
2-1 Hirosawa, Wako-shi
Saitama 351-0198, Japan
http://www.riken.go.jp/

Date Visited: 8 November 1999

WTEC Attendees: A.M. Chaka (report author), R.S. Jones, P.A. Kollman, H. Morishita, P. Vashishta

Hosts: Prof. Toshikazu Ebisuzaki, Director, Advanced Computing Center
Tel: 048-467-9414, Fax: 048-467-4078,
Email: ebisu@postman.riken.go.ip
Dr. Tetsu Narumi, narumi@atlas.riken.go.jp
Dr. Toru Yao, Consultant, Representative from the Genomic Sciences Center
Tel: +81-48-467-9789, Fax: +81-48-467-8149,
Email: yao@riken.go.jp

BACKGROUND

The WTEC Panel met with representatives from two different divisions in RIKEN, the Advanced Computing Center and the Genomic Sciences Center.

RIKEN was founded in 1917 as the first fundamental research institute in Japan. In 1948 the RIKEN Foundation was dissolved by the United States and KAKEN, the Science Research Center Inc., was established. In 1958 the Japanese government reestablished RIKEN as a semi-public corporation, completely supported by the Science and Technology Agency.

RIKEN places a high priority on building links with foreign scientists to help stimulate creativity. Out of 2,500 visiting researchers every year, 500 are from more than 40 foreign countries. In addition, RIKEN sends its researchers to other countries to work on joint projects such as the United Kingdom's Rutherford Appleton Laboratory and Brookhaven in the United States. There are very few permanent research positions. The permanent staff is largely to provide user support in the various labs and centers.

RESEARCH ACTIVITIES

The main site, located in Wako, Saitama, consists of 40 labs in wide ranging areas of physics, chemistry, engineering, biology, and medicine. In addition to the Computational Science Lab, which the panel visited, examples of the other 40 labs are the Cyclotron Lab and labs for applied laser chemistry, magnetic materials, polymer chemistry, organometallic chemistry, microbiology, and regulation of plant functions.

Also located at the main campus are the Frontier Research Program, the Brain Science Institute, and the Genomic Sciences Center. There are six additional research centers located across Japan in Photodynamics, Life Science, Biomimetic Control, Earthquake Research, and the Super Photon-ring 8 GeV synchrotron facility in Harima.

Advanced Computing Center

Areas of research are extremely varied. Specific research projects are decided upon by individual researchers. Hence, there is a lot of freedom for researchers to explore their personal interests.

A few highlights of computational research projects are as follows:

- The Computational Biomechanics Research Group (A. Makinouchi) is developing a simulated human body to enable research in areas such as eye surgery, blood flow in vessels, and the muscle and skeletal system.

- Mineral formation and structure
- Calculating the optical absorption spectrum of hydrogenated Si nanocrystallites containing 4300 atoms
- Homogeneous nucleation in vapor phase—Kenji Yasuka, Keio University
- Image technology (superconducting tunnel junction, high-speed CCD camera, massive photometry observation for astronomy)
- Structure and function of proteins studied by molecular dynamics simulation—Atsushi Suenaga
- MD simulations of the HIV-1 protease substrate complex
- *Ab initio* calculations of the *ras* p21 molecular switching mechanism—Noriyuka Futatsugi
- Order(N) algorithm development (Yasuoka and Matsumoto 1998).

Tetsu Narumi gave a presentation describing MDM, the RIKEN Molecular Dynamics Machine (http://atlas.riken.go.jp/~narumi/MDM/). When the MDM system is completed at the end of 2000, it is expected to achieve 100 Tflops peak and 30 Tflops sustained performance and to be capable of doing half a million time-steps per day on a million atoms. The MDM consists of three integrated computer systems:

- WINE-2:
  - 50 Tflops with 2,688 WINE-2 chips
  - Wavenumber-space part of Coulomb force
- MDGRAPE-2:
  - 50 Tflops with ~3,000 MDGRAPE-2 chips
  - Real-space part of Coulomb and van der Waals forces
- Host:
  - Workstation cluster of 50 Gflops
  - Bonding force, time integration, I/O, etc.

The Host sends coordinates of atoms, then receives calculated forces from WINE and GRAPE. The calculation cost on the specialized subsystem is O(N3/2), while that on the Host and on communication between them is O(N). By the second quarter 2000, they plan to integrate the GRAPE-2 system with a Fujitsu VPP5000 supercomputer. By 2003, the MDM-2 machine is expected to be able to achieve 1 Pflop.

The cost of the system is $5 million. The VPP on-site costs $30 million and runs at 0.3 Tflops peak. The Advanced Computation Center collaborates with IBM researchers, and it is planning to start a small company to market the MDM system. They may begin to market the MDM system through this company by next February or March.

The Advanced Computing Center believes in making this fast machine available to all researchers. So far, it has been very successful.

Genomic Sciences Center

The RIKEN Genomic Sciences Center (GSC) was founded October 1, 1998, to be the central genomic research institute in Japan. Research is currently conducted at three separate facilities until the central location in Yokohama is completed in October 2000, which will provide space for more than 300 researchers. The GSC has received ¥8-10 billion per year from the Japan Science and Technology Agency.

The GSC has three main research programs (Genome Exploration, Human Genome, and Protein) which, taken together, have the stated intent to form a complete, integrated, and systematic approach to the science of life.

- The first objective of the Genome Exploration Research Group is to develop a gene encyclopedia by constructing full-length gene clones and compiling the genetic information. Once compiled it will then

be used for the study of the genetic background of diseases, the functional analysis of proteins, and the cascade regulation of gene expression.

- The Human Genome Research Group is part of an international collaborative effort to sequence the entire human genome.
- The Protein Research Group is focused on determining the three-dimensional structure of proteins by NMR and developing an encyclopedia of the structure of functional domains, called folds, to enable the prediction of the three-dimensional structural form of a protein from its amino acid sequence.

DISCUSSION

Advanced Computing Center

At RIKEN computer simulation is seen as an increasingly important tool to support theoretical and experimental research. Research is being done to develop dedicated hardware and software for high-speed computers specifically for research on the formation and deformation of metallic materials and 3D simulation of protein structures. The 1998 Annual Report (RIKEN 1998) states that "by combining these dedicated computers with general purpose computers, we hope to make massive leaps forward in computational simulation."

Dr. Ebisuzaki has been the Director of the Advanced Computing Center at RIKEN for the past four years. He is an X-ray astronomer by training. Sixteen years ago at the University of Tokyo he initiated a computer effort to do galaxy dynamics and gravitation forces. Gravitational forces, like Coulombic forces, are very long range and hence require significant computer resources. In 1989 his group began to develop a special-purpose computer called GRAPE ("Gravity Pipe" for pipeline) to calculate long-range gravitational forces. He understood that this type of computer could also be used for molecular dynamics (MD) simulations in which other long-range forces such as coulomb forces play a significant role. His group began applying these simulations to proteins nine years ago. When Dr. Ebisuzaki came to RIKEN four years ago, he decided to build a special computer to do MD calculations on proteins.

A big problem is that the software as well as the hardware must be tailored to this unique computing system. They want to be able to easily port FORTRAN codes to their machine and to other machines.

They have a strong connection with researchers from the UK, and hence they started with DL-Poly from Daresbury Laboratories. They collaborate with Thomas Hoover at ANU. At RIKEN they have developed their own MPI (Message Passing Interface) device driver and basic libraries. Using the user-friendly library they have developed, they were able to port GROMOS. They will port AMBER, CHARM and others to their computer system. They have also developed a fast massively parallel FFT code.

The members of the computing center believe there is effective interaction between experimentalists and theorists at RIKEN. One of the best mechanisms is for senior researchers to get together for a beer at 5 pm and discuss their work.

Genomic Sciences Center

Dr. Yao, who has a one-year appointment at the GSC, has been working on computations applied to many chemical as well as biological problems. He initiated a large research group on protein analysis at the Protein Engineering Research Institute in Osaka in 1986. He has been collaborating with panelist Peter Kollman for more than ten years.

The Japanese government funds many activities in computational biology and genomic informatics in addition to the GSC, most notably the JBIC (Japanese Biological Informatics Consortium). This is a large consortium with membership including more than 60 companies and more than 35 research institutions. It was established in Tokyo in November 1998 with a grant from MITI to develop unique software and databases for biological problems. Dr. Yao has worked for the JBIC as an advisor from the beginning.

The goal of JBIC is to transfer technology and most of the research done in academic institutions to industry. What is very unique about this consortium is that it includes both pharmaceutical and computer companies. This hasn't happened before in Japan. Right now the research is done in separate institutions and companies, but it will all come together in one center to be established soon. The government has spent ¥3 billion over 2 years. Companies do not need to invest money to join the consortium.

CONCLUSIONS

The current generation of dedicated molecular-dynamics computer architecture developed at the RIKEN Advanced Computing Center is achieving very high performance.

The JBIC, which brings together pharmaceutical *and* computer companies as well as academic institutions at no cost to the members, is a unique and potentially very effective means of both advancing the state-of-the-art science and transferring it directly to industry.

REFERENCES

RIKEN. 1998. *Annual Report.*

Yasuoka, K., and M. Matsumoto. 1998. *J. Chem. Phys.* 109, 8451.

Site: **Taisho Pharmaceutical Co., Ltd.**
1-403, Yoshino-cho, Ohmiya-shi
Saitama, 330-8530, Japan
http://www.taisho.co.jp/

Date Visited: 9 November 1999

WTEC Attendees: P.A. Kollman (author), A.M. Chaka, R.S. Jones, P. Vashishta, H. Morishita

Hosts: Dr. Kunihiro Kitamura, Director, Ethical Business Strategy Section, (03) 3985-1111, FAX (03) 2985-0716; kitamura@ccm.taisho.co.jp
Dr. Susumu Yamanobe, Molecular Science Laboratory, yamanobe@mol.rd.taisho.co.jp
Dr. Yuji Takaoka, Molecular Science Laboratory, takaoka@mol.rd.taisho.co.jp
Dr. Hajime Takashima, Molecular Science Laboratory, takasima@mol.rd.taisho.co.jp
Dr. Koji Oda, Molecular Science Laboratory, oda@mol.rd.taisho.co.jp
Dr. Hiroo Miyagawa, Molecular Science Laboratory, miya@mol.rd.taisho.co.jp
Dr. Miyoko Yashiro, Molecular Science Laboratory, yashiro@mol.rd.taisho.co.jp

BACKGROUND

Taisho Pharmaceuticals is the leading pharmaceutical company in Japanese over-the-counter drugs (69% of revenues, FY99), but it also has a significant role in prescription drugs (31% of revenues). Its 232 billion yen of sales ($2.1 billion) are 90% based in Japan.

Dr. Kitamura is head of the Molecular Science Laboratory, the molecular modeling group at Taisho. It consists of him and the other six scientists present at the meeting. Kitamura is a physical chemist who was at Osaka University and has been at the company for 15 years. He is now one of the directors, and he does financial modeling for the company as well as leading the molecular modeling group.

The rest of the group has a wide range of backgrounds. Dr. Miyagawa has background in molecular simulations sufficient for him to work on the development of an MD engine (special-purpose hardware for molecular dynamics). Takashima has a quantum mechanics background, and there are thoughts to develop a special-purpose computer for molecular orbital calculations. Oda and Takaoka have been trained in the lab of N. Go at Kyoto University, a top lab for computer simulation of protein systems. Yamanobe has a background in polymer science from the Tokyo University of Agriculture and Technology, and Yashiro was trained in quantum chemistry at Rikkyo University.

R&D ACTIVITIES

Dr. Kitamura began the molecular modeling in the company 13 years ago. His point of view is that when the biology is clear, modeling is crucial. The modeling group at Taisho is unusual in its commitment to significant hardware and software development within the company. The MD engine developed in the group is a collaboration between Taisho, Fuji Xerox, Prof. Amisaki at Tottori University, and Prof. Kusumi at Nagoya University; the second generation of it is about to be completed. The seven scientists in the modeling group provide support to 60 organic/medicinal chemists. The members of the modeling group provide computational support to the synthetic chemists but are also encouraged to develop new software approaches, the feeling being that this should give the company a competitive advantage.

There is one X-ray crystallographer in the company, and they collaborate for structure determination with Prof. Ishida (Osaka University of Pharmaceutical Sciences). There are efforts to work closely with the combinatorial chemistry group, to aid the combinatorial chemists to use Web-based methods in library design. Oda, Takaoka and Yamanobe all work on some aspect of library design, using (Yamanobe) neural networks, decision trees and statistical analyses. It is clear that combinatorial chemistry and its implications

are a focus for modeling at Taisho. In the areas of chemical informatics and bioinformatics, part has been outsourced and part developed in-house.

They do not appear to do extensive modeling of metabolism, distribution and excretion.

In the past five years, 35 molecular biologists have been newly hired, and this expanded effort has led to the identification of 10 proteins as interesting drug targets.

An application of MCSS (Multi Conformation/Simultaneous Search) including solvation was presented at the meeting, and an interest in integral equation methodology was expressed. There are 30-40 targets in the company at any one time and modeling participates in about 20% of these projects, most of which have target structure information. The number of targets with known structure is in the range of 20%; this percentage has been increasing recently. There has not been modeling of G-protein-coupled receptors, but recently, discovery that a new gene product is a G protein may change that situation.

As noted above, Taisho has developed two generations of MD engines for rapid molecular dynamics calculations. They spend about $200,000/year for outside software, including Quanta/Insight2 for visualization, AMBER, Gaussian and MOPAC. The individuals each have SGI workstations for the less demanding calculations and for visualization.

DISCUSSION

It appears that the modelers are appreciated for their contribution to design projects—the group was upbeat regarding what they were contributing to design and combinatorial chemistry. Oda noted that in the area of combinatorial chemistry, 30% success was enough. Although no discussion of credit apportioning was discussed, there appeared to be a commitment in the company to modeling and a belief that this technology, particularly in cases with known structure, has been and would continue to be very valuable.

Finally, a number of future needs were noted by the group:

- Predicting 3D structure of proteins from sequence
- "Virtual reality" interactive MD simulations of proteins in water
- More accurate potentials for molecular mechanics and docking, such that more quantitative predictions could be made
- Better solvation models
- New QM/MM models
- More effective consideration of protein/ligand flexibility in DOCKing or LUDI modeling of ligands in active sites of proteins

Site: **Takeda Chemical Industries Ltd.**
17-85 Juso-honmachi 2-Chome, Yodogawa-ku
Osaka, 532-8686, Japan
http://www.takeda.co.jp/

Date Visited: 10 November 1999

WTEC Attendees: P.A. Kollman (author), A.M. Chaka, R.S. Jones, P. Vashishta, H. Morishita

Host: Dr. Yoshio Yamamoto, Research Head, Discovery Research Laboratories,
(81) 6-6300-6401, Fax (81) 6-6300-6086, Yamamoto_Yoshio@takeda.co.jp

BACKGROUND

Takeda Chemical Industries Ltd. is one of the largest pharmaceutical companies in Japan. In FY99, Takeda had over ¥845 billion ($7.9 billion) in sales, 21% outside Japan. Its sales are dominated by pharmaceutical products (prescription and OTC, 71%) with important contributions of chemical products (13%, primarily polyurethane and elastomers) and food products (9%).

There are eight persons in the modeling/X-ray/NMR group at Takeda. The group leader is Dr. Yamamoto, who got his MS. in crystallography in 1980 and his PhD in biophysics of muscle contraction in 1983. He joined Takeda in 1983 and, as an employee of the company, spent 1986-89 at PERI, the protein engineering research institute, and 1993-94 on sabbatical in the Kuntz group at UCSF. There are two others in the group trained in NMR (one PhD, who has spent time with G. Wagner of Harvard), two X-ray crystallographers, and others with physical organic/synthesis, inorganic quantum chemistry and fiber diffraction of muscle backgrounds.

R&D ACTIVITIES

The group is responsible for both structure and modeling, and some scientists in the group do both structure determination and modeling. One person is responsible for library design. There are five Discovery Research Labs at Takeda. Molecular modeling and structural analysis is located in one of these labs, which also includes general synthesis, automated synthesis, and combinatorial synthesis. There is also a Pharmaceutical Research Division, which includes synthetic chemists. Thus, there are 3 to 5 modelers who support the activities of organic chemists. Of the active pharmaceutical projects at the company, modeling support is provided for 10. Thus, each modeler has 2-3 projects at a given time.

Of the 10 projects, 20% have a detailed three-dimensional structure of the target, 20-30% involve homology model building, 20-30% G-protein-coupled receptor modeling; for 20-30%, no target structure is available. In that case, pharmacophoric modeling must be used. The number of structure-based design projects are increasing. Ten years ago, only 20% of the structures had a target that was known or could be reasonably modeled through homology.

The group has 10 SGI O2s and some PCs. The software includes Insight, Discover, Ludi, Catalyst, DMol and Homology from MSI and combichem software-clustering (Daylight), Sybyl-Diverse Solutions (TRIPOS) and Tsar/Corina (Oxford Molecular) and academic software SCWRL (Cohen, UCSF), DOCK (Kuntz, UCSF), LIGPLOT (University of London) and SCORE (Chinese group).

DISCUSSION

Some of the organic chemists are interested in modeling and others less so. There is a more synergistic interaction with the combinatorial chemistry group, where library design and diversity analysis is done in collaboration with the modelers and with the younger organic chemists.

If the modelers contribute a critical idea, they can be included in a patent. This has happened 2-3 times in Yamamoto's recollection. Many of the organic chemists view modeling as not essential but helpful.

Publication usually occurs in collaboration with synthetic chemists and is not particularly encouraged. However, Yamamoto brought 14 publications by Takeda modelers to our attention, published during the period 1992-present. These publications covered a wide range of areas—modeling of G-protein-coupled receptors, acetylcholinesterase inhibitors, NK1 antagonists, GPIIb/IIIa antagonists, LHRH antagonists, Taxol and analogs, HTLV protease, cathepsin L, and analysis of reaction mechanisms.

Outside interactions are used productively. The group has a collaboration with Dr. Itai's venture company, IMMD (Institute of Medicinal Molecular Design). Takeda has joined projects for using a synchrotron for X-ray structure determination and has considered/is considering supporting university research in key areas. Finally, Takeda has a collaborative effort with Human Genome Sciences and SmithKline Beecham at Tsukuba, where there is also a group in bioinformatics. In the future, more interactions are expected with this group.

CHALLENGES CURRENTLY AND IN THE FUTURE

There is a great need for more accurate structures, and thus the need for more protein crystallography is greater than the need for more modeling and modelers. More powerful hardware is needed for more rapid DOCKing and better software for calculating interaction free energies of ligands with targets. More generally, judging by the standard of the large impact that XPLOR has had on crystallographic computation, the impact of computational modeling has been disappointingly small to Yamamoto. Until one can really predict a much higher percentage of hits or a much improved activity, the impact of modeling on pharmaceutical design will be disappointingly limited.

Site: **Toshiba Corporation, Research & Development Center**
1, Komukai Toshiba-cho, Saiwai-ku
Kawasaki-shi 212-8582, Japan
81-44-549-2056
http://www.toshiba.com/

Date Visited: 11 November 1999

WTEC Attendees: E.B. Stechel (report author), P.T. Cummings, S.C. Glotzer, R. Miranda, K. Morokuma, P.R. Westmoreland

Host: Yasuo Ikawa, Group Manager, Strategic Planning Group, Research Planning Office, (81) 44-549-2053, yasuo.ikawa@toshiba.co.jp
Hiromi Takahara, Research Administration Group, Research Planning Office
Yoshinori Fujimori, Chief Specialist, Research Administration Group, Research Planning Office, fujimori@rds.rdc.toshiba.co.jp
Shinobu Fujita, Chief Specialist, Strategic Planning Group, shinobu.fujita@toshiba.co.jp
Francis Minoru Saba, Specialist, Strategic Planning Group, minoru.saba@toshiba.co.jp
Koichi Kato, Chief Research Scientist, Advanced Materials & Devices Laboratory, kato@arl.rdc.toshiba.co.jp
Shigenori Tanaka, Senior Research Scientist, Advanced Materials & Devices Laboratory, (81) 44-549-2299, tanaka2@arl.rdc.toshiba.co.jp
Satoshi Itoh, Senior Research Scientist, Advanced Materials & Devices Laboratory, satoshi.itoh@toshiba.co.jp
Yuusuke Sato, Research Scientist, Mechanical Systems Laboratory, ysato@msl.rdc.toshiba.co.jp
Shinji Onga, Chief Research Scientist, Advanced LSI Technology Laboratory, onga@amc.toshiba.co.jp

BACKGROUND

The origins of Toshiba Corporation date back to the founding of two independent companies: "*Tanaka Seizo-sho*" in 1875 (renamed "Shibaura Engineering Works Co. Ltd" in 1904) and "*Hakunetsu-sha*" in 1890 (renamed "Tokyo Electric Company" in 1899). Those two companies merged in 1939 to become "Tokyo Shibaura Electric Co. Ltd.," which was subsequently renamed Toshiba Corporation in 1978.

Toshiba Corporation manufactures products covering information and communication systems (~41% of revenues), information equipment and consumer electronics (19%), power systems and industrial equipment (~18%), and electronic components and materials (~20%). The remaining ~2% is services and other, such as finance, insurance, real estate and logistics. By region, Toshiba sales are 15.9% North America, 60.1% Japan, 11% Asia, 10.6% Europe and 2.4% other. The Toshiba Group (consolidated base) had 198,000 employees and net sales of 5,300 billion yen ($47 billion) for the fiscal year ending March 31, 1999.

On April 1, 1999, Toshiba Corporation reorganized its 15 business groups into eight in-house companies and one joint venture:

- Information and Industrial Systems & Services Company
- Digital Media Equipment & Services Company
- Power Systems & Services Company
- Semiconductor Company
- Display Devices & Components Company
- Medical Systems Company
- Home Appliances Company
- Elevator and Building Systems Company

Separate from the in-house companies, headquarters activities are divided into Corporate (administrative) Staff, Corporate Projects, and Corporate Support. The Corporate R&D Center is within the Support Services segment.

Total research is 6% of sales, or about $2.6 billion annually. Basic research (longer than five year time horizon) is conducted at the Corporate R&D Center. The center has 1,300 employees (150 PhD and 700 MS). Research activities include display materials, Li_2ZrO_3 lithium ion batteries, CO_2-absorbing material, imaging, memory storage, and environmental technologies.

R&D ACTIVITIES

Toshiba research and development does not have a modeling group. It does have modelers integrated into several laboratories, such as Advanced Materials and Devices Laboratory, Mechanical Systems Laboratory and Advanced LSI Technology Laboratory. The LSI group was established in 1987. It has two modelers and close interactions with the engineers. They have developed their own sophisticated codes for silicon using three-body covalent models for the force field. After developing the force field, they also developed their own SiO_2 molecular dynamics code. They are starting to use DFT and tight-binding models as well.

Examples of molecular modeling projects include the following:

- *Atomistic Simulation of Si Oxidation (K. Kato).* Electronic density-functional calculations (Perdew 91) are being used to investigate O_2 dissociative attachment on Si and the energy barriers for layer by layer oxidation. On Si(100), there is a barrierless adsorption onto the top layer, but there are questions about whether the oxygen adsorbs molecularly or dissociatively and what the barriers are for the different layers. The differing barriers are best explained by curve-crossing between triplet and singlet surfaces, supported by experiments of Watanabe at NEC. This work is now being extended by plane-wave pseudopotential calculations using a new JRCAT code. Kato began this work at JRCAT and joined Toshiba two years ago.
- *Dendrimers: Simulations of Supramolecular Materials (S. Tanaka).* Dendrimers have possible applications as photofunctional materials, in molecular recognition and in catalysts. Calculations are mostly semi-empirical MO (MNDO, AM1, PM3) for structural properties, vibrational properties, ground and excited-state electronic structure, excitation and relaxation dynamics, and chemical reactivity. This work is exploratory and does not yet have experimental collaborators.
- *Reaction Analysis Tools for Engineers (Y. Sato).* This work is intended to provide useful integrated reaction-engineering tools for the CVD process engineers. Calculations are made for reaction rate constants, diffusion modeling, mechanisms, and qualitative guides to controlling behaviors like resistivity, grain size, and roughness. Elementary reactions are automatically shown from the semi-empirical level DRC (dynamic reaction coordinate) calculation, in which source molecules with the kinetic energy move based on the potential energy calculated by the semi-empirical level in MOPAC. The results are viewed with Re_View and XMol. The accurate activation energy is determined at MP2/6-31G(d,p) level using Gaussian and SPARTAN. Semi-empirical level slab calculations or the classical MD calculations where potentials are suited for reaction are future challenges. It also has proven important to set up data structures for reaction types, aiding the practical organization of the results. In this example, the engineering scientist is proactively anticipating the need, rather than being driven by requests from process engineers.
- *TCAD and its Application to LSI Design Phase (S. Onga).* Researchers have demonstrated hierarchical TCAD tools that include a molecular-level simulator. The TCAD tools span the range from first-principles atomistic calculations to process design. For example, a dislocation generator was created by incorporating stresses in an NPD ensemble calculation. A particular goal is to incorporate tight-binding molecular dynamics into TCAD. Experimentalists provide validating data and the strong motivation to develop the tools. (Dr. Onga recently spent two years at Stanford University.)

From the perspective of Toshiba researchers, breakthroughs that could have large impact include both the likely advances in computer science and density functional advances for excited states, finite temperature, external electric fields, and other applications.

CONCLUSIONS

The role of research at the Toshiba Research & Development Center is to produce patents, develop new concepts and to improve existing products and processes. Much of the modeling work in the materials department is very exploratory and very basic. The work in mechanical systems is proactive and designed to deliver a useful tool to the engineers. The work in the LSI department is highly integrated with experiment and the end user.

REFERENCES

Hiraoka, Y.S. 1997. Ab initio molecular orbital study on the As-stabilized GaAs(001)-(2x4)b1 surface. *Surface Science* 394, 71-78.

Schimizu, T., and Kawakubo, T. 1998. First-principles approach to the effect of c-axis elongation of $BaTiO_3$ thin films. *Japanese Journal of Applied Physics* 37, 235-237.

Tada, T., and R. Yoshimura. 1996. Theoretical studies on hydroxylation mechanism of hydrogen-terminated Si surface in aqueous solutions. *Physics Letters* A 220, 224-230.

Watanabe, H., et al. 1998. Kinetics of initial layer-by-layer oxidation of Si(001) surfaces. *Physical Review Letters* 80, 345-348.

Site: **Toyota Central Research & Development Laboratories, Inc.**
http://www.tytlabs.co.jp/

Reporter: P.R. Westmoreland

Contact: Dr. Atsuo Fukumoto (wien@cmp.tytlabs.co.jp), Senior Researcher, Applied Mathematics and Physics Lab, System Engineering Division I

Toyota Central Research & Development Laboratories is a company jointly owned by members of the Toyota Group. Best-known of the companies in the Toyota Group is Toyota Motor Corporation (http://www.global.toyota.com/). The latter is one of the world's largest companies, ranking third in volume of motor vehicles (FY98: GM 8.1 million units, Ford 6.8 million, Toyota 5.2 million) and fourth in sales (GM $160 billion, DaimlerChrysler $155 billion, Ford $143 billion, Toyota $88 billion).

Toyota CRDL was formed in 1960. Its 956 employees presently occupy a modern site just east of Nagoya. The company's research includes work requested by companies in the Toyota Group as well as proprietary work that it initiates. The laboratory's technical departments include System Engineering and Electronics, Mechanical Engineering (including human-factors engineering), Materials (including catalysis, batteries and fuel cells, and biotechnology), Research Fundamentals Technology, and Frontier Research. Its activities have won twenty-three *R&D 100* Awards.

Computer modeling of materials began at the company in 1988, focusing on semiconductors for high-temperature and high-power electronic devices (see references below). Since 1997, it has re-focused on materials related to environmental problems. Four people carry out the principal responsibilities, using both vector supercomputers and workstations. Most calculations use first-principles pseudopotential methods implemented in codes written at Toyota CRDL.From their perspective, a key need for the field is a reliable method or set of methods for complicated structures. Almost as important is the need to evaluate and assess the limitations of existing methods. As one way of working toward these objectives, they are members of the CAMP project (General Purpose Computational Material Physics Platform), a software development project of 10 companies partially supported by the IPA (Information-technology Promotion Agency, MITI). That project is described in more detail in the NEC report.

At the same time, they see a number of promising technical developments. Chief among them is the ability to calculate accurate free energies (materials properties at finite temperature) and the advent of transferable interatomic potential functions.

REFERENCES

Fukumoto, A. 1990. First-principles pseudopotential calculations of the elastic properties of diamond, Si, and Ge. *Phys. Rev. B* 42, 7462.

Fukumoto, A. 1996. First-principles calculations of p-type impurities in cubic SiC. *Phys. Rev. B* 53, 4458.

Fukumoto, A. 1997. First-principles calculations of impurity states in 3C-SiC. *Phys. Stat. Sol. (B)* 202, 125.

Fukumoto, A., and K. Miwa. 1997. Prediction of hexagonal Ta_2O_5 structure by first-principles calculations. *Phys. Rev. B* 55, 11155.

Miwa, K., and A. Fukumoto. 1993. First-principles calculations of the structural, electronic, and vibrational properties of gallium nitride and aluminum nitride. *Phys. Rev. B* 48, 7897.

Miwa, K., and A. Fukumoto. 1995. Theoretical study of Ti adsorption on Si(001) surfaces. *Phys. Rev. B* 52, 14748.

Sites: **Short reports from other Japanese companies:**
Fuji Photo Film Co., Ltd.
Japan Polyolefins, Ltd. (Showa Denko K.K.)
JSR Corporation
Mitsui Chemicals
Nippon Zeon Co., Ltd.
Sumitomo Chemical Co., Ltd.
UBE Industries, Ltd.

Reporter: P.R. Westmoreland

It was not possible to visit all the companies of interest. Information was collected for several Japanese companies from the public domain (company Web sites, public presentations) and by comments of company researchers. Opinions and speculation shown below are those of individuals, not necessarily of the companies.

FUJI PHOTO FILM CO., LTD.

Fuji Photo Film (http://home.fujifilm.com/) is a major producer of photographic film, digital imaging equipment, and supplies for these end users. Sales in FY1999 were 808 billion yen (about $8 billion).

Molecular modeling activities began about 15 years ago, aimed toward applications into materials for photography. Current work involves several people working on properties and kinetics of functional dyes, developing reagents, and silver halide. One of Fuji's staff members, Hiroo Fukunaga (fukunaga@ashiken.fujifilm.co.jp), is a participant in the Doi Project for modeling polymers (described separately).

From Fuji's perspective, the biggest needs are abilities to handle solvent effects and defects in solids, plus the need for faster CPUs. Both molecular design and reaction design are promising applications for the future.

JAPAN POLYOLEFINS, LTD.

Japan Polyolefins, Ltd., is a component of Showa Denko K.K. (http://www.sdk.co.jp/), a large Japanese chemical company involved in petrochemicals, chemicals, electronics, and inorganic materials, most significantly aluminum. Net sales for Showa Denko in FY99 were ¥663 billion (nearly $6 billion), of which about 30% was from the petrochemicals business that includes Japan Polyolefins. Its polypropylene business was split off in June 1999 into a new joint venture, Montell SDK Sunrise, Ltd. As a result, Japan Polyolefins is completely in polyethylenes.

Its modeling is focused on polymer properties and synthesis, including metallocene catalysis. The company has had people working on molecular and materials modeling for about ten years. Presently there are five people in computer-aided engineering and three in molecular modeling. Because predictions of rheology and other physical properties are longer-term goals, it is appropriate that they participate in the Doi Project. Hiroyasu Tasaki (tasaki@zoom.cse.nagoya-u.ac.jp) is their representative, partway through his three-year assignment to the project.

JSR CORPORATION

JSR Corporation (http://www.jsr.co.jp/) is a producer of synthetic rubber, plastics, and fine chemicals. It was Japan Synthetic Rubber, Ltd., for its first 40 years until renamed in 1997. Its 1998 sale revenues were ¥163 billion (about $1.4 billion), 48% from its synthetic rubber business.

Modeling is focused on metallocene synthesis and properties of polymers. The company's three people in molecular modeling represent ten years of work in the field. They see multi-scale modeling of polymer structures as an important need for them to achieve true molecular design. They take part in the Doi Project.

MITSUI CHEMICAL

Mitsui Chemical is a component of Mitsui & Co., Ltd. (http://www.mitsui.co.jp/), a diversified conglomerate of 900 companies. Mitsui reports that of its ¥14 trillion of trading transactions ($130 billion, FY99), chemicals represented ¥1,831 billion ($17 billion). Mitsui Chemical has three major sectors: petrochemicals and polymers, specialty chemicals and plastics, and fertilizers and inorganic chemicals.

It has been active in molecular modeling for 15 years, with interests including polymers, catalyst design, materials development, biochemicals, and process design. Fifteen people work in the area, including Dr. Takashi Igarashi, who is assigned through the Japan Chemical Innovation Institute as Senior Research Manager of the Doi Project. Accordingly, large-scale simulation is a concern that they are addressing. Long-term, they see the larger target of polymer design (reaction, rheology, and computer-aided engineering) as especially promising.

NIPPON ZEON CO., LTD.

Nippon Zeon Co., Ltd. (http://www.zeon.co.jp/) is a mid-sized chemical company centered around technologies in synthetic rubber and other polymers. Employing about 2600 people, it had FY98 sales of ¥122 billion ($1.1 billion).

It has been involved in molecular modeling for about 10 years, presently involving two specialists. Polymer properties and catalytic synthesis are both of interest. Meso-scale simulation is an important need for them, and consequently they participate in the Doi Project at Nagoya. Long-term, they see catalyst design as especially promising.

In late 2000, Mitsui and Sumitomo announced that they will undergo a 50-50 merger, to be completed in October 2003. Chemical sales of $16 billion would result from a combination of the two companies. They intend to merge their polyolefin businesses by October 2001 (C&EN 2000).

SUMITOMO CHEMICAL CO., LTD.

Annual sales of ¥928 billion ($7.7 billion, FY98) make Sumitomo Chemical Co., Ltd. (http://www.sumitomo-chem.co.jp/) one of the largest chemical companies in Japan. The company is organized into five business sectors—Basic Chemicals, Petrochemicals, Plastics, Fine Chemicals, and Agricultural Chemicals—but it also has a range of subsidiaries including Sumitomo Pharmaceuticals (http://www.sumitomopharm.com). It has ten research laboratories in Japan, half of which do sector R&D while the other half work on crosscutting interests or long-term topics.

Its 20 to 30 molecular modelers give it one of the largest industrial groups in this field in the world. (The number at any time varies depending on need.) The original efforts in 1982 were in agricultural chemicals, using ACACS solvers for molecular design. Since then, modelers have carried out wide-ranging work in chemicals, life sciences, and materials: agro-chemicals (pesticide design), functional dyes, polymers (including liquid crystals), catalysts (Ziegler-type), photo-electric materials, and synthesis routes for pharmaceuticals. An example of code and method development is their work on a kinetic lattice model for Monte Carlo dynamics of chemical systems (Mol. Sim. 1999).

Improvements in hardware and software are watched closely. Presently, the hardware is a mixture of engineering workstations and parallel machines. Much faster CPUs and more memory will continue to greatly increase the number and complexity of problems that can be addressed. So will the increased use of cheap PC clusters, which can be faster than an expensive workstation.

Both in-house and commercial software are used for computation. New computational software should be accompanied by simple and flexible interfaces, important in industry because research themes vary so much. Operating systems and system software should be improved as well. Writing parallel code is harder than it should be because automatic parallelizers do not work well—or sometimes at all.

Precision frequently is not the most important thing in successful industrial use of modeling. "Simple and fast" is often much more important to practical work. The researcher in industry needs a sense of the key properties of the target material.

Nevertheless, they see combining chemistry and physics as still being a challenge for materials modeling. A specific current limitation is the need to adapt data-mining methods for materials design. Expected developments that will help them greatly include mesoscale simulation, design of synthesis routes, and improved molecular design. As a measure of the importance of these topics, the late Dr. Mototsuga Yoshida of Sumitomo was assigned to the Japan Chemical Innovation Institute and served in the Doi Project as a subleader.

Likewise, Sumitomo has participated actively in the CAMM and CAMP projects (http://www.camp.or.jp/) through Dr. Yasunari Zempo (zempo@tuc.sumitomo-chem.co.jp) and Dr. M. Ishida of the Tsukuba Research Laboratory. By focusing on the calculations and the physics/chemistry of phenomena while avoiding their confidential areas of work, members of these projects have developed a mutually useful code for *ab initio* molecular dynamics, CAMP-Atami (see the report from NEC).

In late 2000, Mitsui and Sumitomo announced that they will undergo a 50-50 merger, to be completed in October 2003.

UBE INDUSTRIES, LTD.

UBE Industries, Ltd. (http://www.ube-ind.co.jp/) is a broadly based energy, chemicals, and materials company. Its chemicals and plastics business has five sectors: basic chemicals, fine chemical, specialty chemicals, plastics, and synthetic rubber. Annual sales for all of Ube Industries was ¥538 billion ($3.1 billion) in FY98, while for chemicals and plastics, sales were ¥210 billion ($1.2 billion).

Ube's molecular and materials modeling interests have mainly been in the areas of polymers and polymerization catalysts, especially catalyst activity. Two people are company specialists in the field, where the company has had activity for the last 10 years. One person, Tatsuya Shoji (shojita@zoom.cse.nagoya-u.ac.jp) is assigned to the Doi Project because large-scale simulation is such a need for the company's applications. Looking ahead, company researchers see rheology and molecular design of new materials as areas where valuable strides will be made.

REFERENCES

Chem. and Engineering News (C&EN). 2000. Nov. 27, p. 6.

Molecular Simulation (Mol. Sim.). 1999. 21, pp. 343-355.

APPENDIX B.3. SITE REPORTS—UNITED STATES

Site: **3M, Inc.**
3M Center
St. Paul, MN 55144-1000
http://www.mmm.com/

Date Visited: 6 December 1999

WTEC Attendees: P.R. Westmoreland (report author), P.T. Cummings, R. Miranda, E.B. Stechel

Hosts:
Frank J. Armatis, Technical Manager, 3M Advanced Materials Technology Center, fjarmatis@mmm.com
Cristina U. Thomas, Group Leader, Applied Material and Modeling Group, 3M Advanced Materials Technology Center, cuthomas@mmm.com
Shih-Hung Chou, Applied Material and Modeling Group, 3M Advanced Materials Technology Center, schou@mmm.com
Richard B. Ross, Applied Material and Modeling Group, 3M Advanced Materials Technology Center (by telephone)
Robert S. Davidson, 3M Corporate Process Technology Center, rsdavidsonl@mmm.com
David K. Misemer, Software, Electronic, and Mechanical Systems Laboratory, dkmisemer@mmm.com
Mark Rustad, Laboratory Information Technology

BACKGROUND

3M Inc. (formerly Minnesota Mining and Manufacturing) is a diversified global materials company with interests divided into six market areas: Consumer and Office Markets, Industrial Products, Electronic and Commercial, Health Care, and Special Materials. Many of these areas have overlapping technical interests. A large portion of 3M's products are sold to other companies, but it also produces well-known consumer products like Scotch tape, Scotchgard fabric treatment, Scotchbrite scouring pads, Post-It notes, and Thinsulate clothing insulation. 3M operates in 60 countries with nearly 74,000 employees, over 38,000 of whom are in the United States. 1998 sales were $15 billion, and $1.02 billion was invested in R&D.

Dr. Armatis described research activities as being currently divided among corporate centers and market technical centers, which focus on the individual market areas listed above. Because the company relies so heavily on innovation, corporate policy is to maintain an "innovation network," where 15% of research activity is set aside for work on ideas that would significantly change the company's businesses.

MOLECULAR MODELING R&D ACTIVITIES

Molecular modeling began in the early 1980s, but the current activities began in 1988 with three people (including Dr. Chou) doing semi-empirical quantum chemistry (MOPAC) and molecular mechanics. At that point, part of the reason was recognition that other people were making use of these methods and that 3M must keep up. Dr. Misemer (whose background was condensed-matter physics and DFT) joined this group in 1991, and Dr. Thomas (a PhD chemical engineer and polymer scientist) joined shortly afterwards. Dr. Ross (quantum chemistry background) came from PPG in 1998.

In the past, molecular modeling work was rather centralized, while now it is distributed. The range of presenters emphasized this distribution. The hosts estimated that for all types of modeling, there are probably 60 specialists company-wide, mostly PhDs working in 6 to 12-person groups within the technical centers. (Dr. Misemer emphasized the need for such large clusters of modelers. By experience, if only one or two

modelers are involved, they spend most of their time having to do things other than modeling, dissipating the potential impact of their modeling abilities.) Most of 3M's modeling is by engineers with a small number of computational chemists and physicists.

Dr. Thomas, group leader of the Applied Material and Modeling Group, described their work as being a corporate-level activity but closely involved with individual market areas. Materials, products, and processes are subjects. Most of their modeling is need-driven. Increasingly, there are requests for modeling at the initial stages, where chances are much better. They are involved in the complete spectrum of activities from electronic-structure calculations to finite-element continuum modeling, although mesoscale modeling is just beginning.

A wide range of successful applications was described, using both quantum chemistry and molecular simulations. Noteworthy aspects are the interactions with experimental chemists and engineers and use in direct computational screening of candidate materials:

- *Design of environmentally friendly hydrofluoroethers* (with Scott Thomas and John Owens of 3M). The goal was to identify solvents for precision and electronics cleaning that would also have atmospheric lifetimes less than 20 years, replacing CFCs. From the methods described by Cooper et al., it was established that atmospheric lifetimes of fluorocarbons were correlated with HOMO energies. With some modeling work done by Scott Thomas, new products were commercialized, and in addition, 11 scientists who worked on developing this product at 3M received a 1997 ACS "Heroes of Chemistry for Chemistry and the Environment" award.
- *Finding a molecule for monitoring potassium in the blood by fluorescence*, to be used during medical procedures involving heart-lung machines. When stimulated by an light-emitting diode, the molecule had to absorb light at wavelengths >380 nm and emit fluorescent light with wavelengths <500 nm. The intensity of emitted fluorescent light is correlated to the concentration of K+ in a patient's blood. Quantum mechanics was used to determine the conformation, and ZINDO was used to predict the absorption and fluorescence wavelengths for novel molecules. All screened molecules are derivatives of 6,7-O,O-[2.2.2]cryptando-4-ethylcoumarin, and a broad patent based on modeling results was issued in 1999.
- *QSAR evaluation of additives for imaging films*, primarily useful for broadening patent coverage and for defensive analysis against patent infringement.
- *Design of a true-yellow fluorescent pigment* (as opposed to yellow-green) for highway signs. Ninety possible derivatives of the existing pigment were evaluated, yielding ten candidates. Five were synthesized, and two of these were true yellow, leading to a ROI (record of invention) filing.
- *Analysis of binding of polymers and molecules on polyethylene.* Dr. Ross described how molecular dynamics codes (MSI) were used to relax a polyethylene surface for 450 ps, and a nonproprietary example of surface binding was then shown with a polyurethane oligomer bound to the surface. In applied studies at 3M, molecule/surface binding has been analyzed to increase understanding of surface adhesion (presented at the 1999 Fall National American Chemical Society Meeting, New Orleans, LA, August 1999). Similar example calculations of polypropylene-polyethylene and water-polyethylene interactions were shown to illustrate new cutting-edge research applications being pursued for interfacial interaction.
- *Development of low-shrinkage polymers for dental applications.* In-house and MSI reactive-molecular dynamics codes was used to examine cross-linking reactions and the topology of the resulting networks (Doherty et al. 1998). The results were used to help decision-making for experimental product development.
- *Polymer spectroscopy and other properties*, calculated to provide full characterization for patents and applications.
- *Catalysis modeling* is carried out in a mostly experimental group, linked also to ASPEN process modeling activities for cost evaluation. Dr. Davidson described efforts to design catalyst structures that will produce desired product structures.
- *Development of guided-wave optics and fiber-optic piping of sensor-molecule photons.*

Internal computing resources include a 16-processor SGI Origin, although they expect to move toward Windows NT as software becomes available. Much computing is done on fast dual-processor workstations, and supercomputer time is rented when necessary. Distributed parallel computing (like Beowulf clusters) is not yet sufficiently standardized for it to be a justifiable choice. MSI software is used heavily, but they also use Unichem, Gaussian, and Spartan (Wavefunction, Inc.).

Collaborations are used when appropriate. Productive activities were cited with Imperial College, the University of Michigan, Los Alamos Scientific Laboratories, and the Material Sciences Consortium of Goddard at Caltech.

DISCUSSION

3M envisions itself as a "new-product engine." This is borne out by the statistic that 30% of sales are derived from products introduced in the last four years. The products are not always the results of new technology, but also of new combinations of technologies. At all technical and managerial levels, there is growing appreciation that modeling can improve the results, reduce the time of product development, and drive innovation. The multitude of technologies and products makes the task of modeling especially challenging since one needs to be able to apply modeling across a diverse set of conditions. It becomes even more important given the large number of products—over 50,000 —based on a wide range of core technologies.

Driving forces for modeling include both productivity and innovation:

- Improved decision-making—analytically driven experiment and design
- Improved positions in intellectual property, both for broader claims and for differentiation
- Improved process optimization and control
- Increasing requirements of customers that 3M deliver performance models along with the physical product

Modelers must recognize how to be part of a systems approach to development, and software companies must recognize what the modelers need. Almost all materials are multi-component, sometimes with properties dominated by additives present in small amounts. Idealized cases often have great value, but they must cope with reality as well. Modeling of solvation is crucial, both for environmental fate and for describing mixture non-idealities.

Overcoming cultural barriers between disciplines and sectors is just as important. Failure of clients to take advantage of potentially useful modeling may be due to skepticism or simply to lack of experience. New hires are often most comfortable with working with modelers, as many have experienced molecular-scale thinking and modeling as routine parts of their education. Regardless, modelers have a double responsibility at 3M: they must be able to understand the language of product development and to communicate their results free of the jargon and technical details of the modeling activity. Validation of a model must be seen as progress toward an answer, not as a detour. That can be hard to understand if it requires some experimental syntheses of poor material candidates. On the other hand, the engineering community can deal much more comfortably with low-precision answers that are known to be accurate enough to be adequate for decision-making.

Looking to the future, QSAR and QSPR analyses will be used more, both for statistical modeling and for designed experiments. More response functions will be developed to be incorporated into finite-element methods. Mesoscale modeling seems promising. At the same time, combinatorial materials science and rapid-screening methods are of growing interest to researchers. Informatics will play an increasing role. Smaller physical structures will be a challenge, as in nanotechnology and MEMS development.

CONCLUSIONS

3M is noteworthy both for bridging the chemical process industries and the materials industries and for its intense focus on innovation. Its recent molecular modeling activity has proven especially effective, making

the modeling group an accepted contributor to product development. Reasons for this success appear to be the attention to doing the right modeling scientifically while integrating specialist activities into the pre-existing development process.

REFERENCE

Doherty, D.C., B.N. Homes, P. Leung, and R.B. Ross. 1998. Polymerization molecular dynamics simulations. I. Cross-linked atomistic models for poly(methacrylate) networks. *Comp. Theoretical Polymer Science,* 8(1/2), 169-178.

Site: **Air Products and Chemicals, Inc.**
7201 Hamilton Boulevard
Allentown, PA 18195-1501
(610) 481-4911; Fax: (610) 481-5900
http://www.airproducts.com/

Date Visited: 18 January 2000

WTEC Attendees: M. Neurock (report author), P.A. Kollman, R. Miranda, K. Morokuma, E.B. Stechel, P. Vashishta, and P.R. Westmoreland

Hosts: Dr. Brian K. Peterson, Engineering & Scientific Modeling, petersbk@apci.com
Dr. Hansong Cheng, Senior Research Chemist, Engineering & Scientific Modeling, chengh@apci.com
Dr. Carlos A. Valenzuela, Manager, Engineering & Scientific Modeling
Dr. John B. Pfeiffer, Director, Research and Engineering Systems

BACKGROUND

Air Products and Chemicals, Inc. is a multinational company with over 17,400 employees worldwide in over 30 countries. They are a leading supplier of industrial gases and related equipment along with both intermediate and specialty chemicals.

Industrial gases and equipment account for approximately two thirds of the company's present business, with sales in N_2, O_2, Ar, H_2, and CO, as well as a variety of fluorocarbon and other specialty gases. Processing equipment makes up a critical component of the industrial gases business. Along with the gases, Air Products markets the optimal equipment and application technology. In terms of intermediate and specialty chemicals, Air Products is divided into vinyl polymers, alkyl and specialty amines, polyurethane intermediates and additives, epoxy curatives, surfactants, and coating additives.

Allentown is the home for most R&D activities. R&D efforts emphasize the following technologies: separation science, polymer science, specialty organic chemicals, inorganic materials, environment and energy, application technology, analytical science, and technical computing. Within Research and Engineering Systems is Engineering & Scientific Modeling, where molecular modeling is located. Its larger roles are as follows:

- Apply and support use of technical software
- Internal consulting on computing technology
- Training and education
- Develop new value-added applications of technical computing
- Identify strategic directions for technical computing
- Assure appropriate MIS infrastructure planning
- Recruit, develop, and motivate technical staff

The technological capabilities of this group cover a diverse set of areas including applied mathematics, engineering and scientific modeling (includes molecular modeling and computational fluid dynamics), technical information and knowledge management, modeling and simulation, and software development and computing infrastructure.

Computational chemistry modeling is the focus only of the two primary contacts, Dr. Hansong Cheng and Dr. Brian Peterson, but they are involved in activities that cover a wide range of businesses and technologies. In addition, they have trained other individuals who have used molecular modeling to some limited extent.

MOLECULAR MODELING R&D ACTIVITIES

Molecular modeling formally started at Air Products over 15 years ago. Initial efforts provided mainly "pretty pictures," which unfortunately convinced management that molecular modeling would play a major role in designing new materials in just a few years.

There was some limited initial success in understanding polyurethane catalysis using simple MOPAC calculations. The results initially appeared to show the correct qualitative behavior. The results were later shown to be quantitatively incorrect, but they still provided initial guidance.

From 1989 to 1992 computational quantum chemical efforts could be described as the "MOPAC Era." Dr. Paul Mathias (a chemical engineer and thermodynamicist now at Aspen Technology) led a number of activities to increase the use of these and other methods, mostly using Biosym software. From the late 1980s to the early 1990s, they tried to build molecular modeling expertise in the company by involving newly hired modelers through the company's new-employee rotational program. Unfortunately, once these young engineers rotated onward, they did not continue to use their computational chemistry skills, but moved on into other engineering or scientific disciplines.

In 1991, Dr. Hansong Cheng was hired and in 1996, Dr. Brian Peterson. Dr. Cheng is primarily in charge of quantum chemical projects. He is a chemist who had come from doctoral studies with Herschel Rabitz of Princeton. Dr. Peterson is primarily in charge of molecular simulation. He is a chemical engineer who came to Air Products after seven years of modeling work at Mobil. The two work closely together in an effort to guide molecular modeling at Air Products. While the group works to help educate others, there are only a few other people using molecular modeling themselves, all on a part-time basis.

Following the addition of Dr. Cheng, the quantum chemical efforts emphasized *ab initio* calculations more heavily. A fair number of their projects require more accurate analyses and therefore lend themselves to *ab initio* calculations. There is a mixture of both short turnaround projects as well as long-term efforts. The first group includes structure-property calculations for selective fluorination and nitration, thermochemical calculations, chemical vapor deposition, and homogeneous catalysis. Longer-term efforts include calculations on gas adsorption and storage, chemical corrosion, and heterogeneous catalysis.

WTEC's hosts described the history of Air Products' work in the area of molecular simulation, including work in the 1980s and 1990s on molecular mechanics, molecular dynamics and Monte Carlo simulation and software (see Table B3.1). Much of the earlier work was spent on small molecules, polymers, liquid surfaces, and force-field developments. It was necessary to build up core modeling strengths and demonstrate credibility for modeling. Later efforts expanded and improved upon these areas, whereby both the size and the complexity of the problems were increased. Current efforts are aimed at examining the interaction of small molecules with pore walls, finding the location of cations in zeolites, and predicting adsorption isotherms.

Hardware and Software

The hardware and software available for molecular modeling has steadily increased. In 1994 the group began using a series of IBM RISC workstations to carry out computational chemistry as well as engineering calculations. Very recently, Research and Engineering Systems purchased a 16-node SGI Origin 2000 along with two dual-node SGI Octane workstations for the computational chemistry effort. Current software that is used includes Spartan (Wavefunction), Jaguar (Schrödinger), Gaussian, MSI codes, Gulp, VASP, as well as algorithms for molecular simulation that were written in-house.

Table B3.1
History of Molecular Simulations at Air Products

Time Period	M. Mechanics Size / Shape	Molecular Dynamics	Monte Carlo	Software
late 80s		discussions		
early 90s		Liquid surfaces	FF development for adsorption	Self-written
mid 90s	Small molecule & polymer QSPR Inorg. structures	Flow through membranes	RAP structure solution Si/Al distribution in zeolites	Biosym scripts + docking Biosym MSI Gulp
late 90s	Small molecule interactions with pores	Polymer structure Liquids Liquid surfaces	Cation location in zeolites Adsorption sims	Self-written NPT MD Free Volume Cation Locator NVT MD

Examples

Design of Fluorinating Agents

The group was highly successful in developing a structure-property database that is used to rapidly predict the most appropriate fluorinating agent for customers' compounds. By carrying out a series of *ab initio* and semi-empirical MOPAC calculations, they established the general features of the reaction mechanism for fluorinating aromatic intermediates. They were able to rank the substrates and reagents in terms of increasing the reactivity of the substrate and increasing the fluorinating agent's reactivity. Their predictions matched remarkably well with what was observed in experiments.

The calculations also helped sales of their fluorinating reagents to their customers because they were then able to establish a simple *ab initio* calculation test that they can carry out to indicate the reactivity of their customer's compounds towards fluorination. In addition, the test also provides a tool for screening the safety of the reaction that would need to be performed.

This tool has had an important impact on the Air Products' business. It ultimately helped to give Air Products an edge in the market place and was a major accomplishment for the molecular modeling group. Expansion of the market depended critically on modeling, which created real customer satisfaction. This was really a first for the company, in that molecular modeling was involved in aiding chemical sales.

Metal Dusting

A critical problem in the synthesis of higher hydrocarbons from CO and H_2 is that of metal dusting, which involves the pitting and decomposition of the metal surface. Air Products spends heavily each year to replace syngas reactors and piping. Over time, the metals are converted into metal oxide and carbides. Initial evidence suggested that the breakdown of the oxide might be responsible for metal dusting. DFT cluster calculations were used to show that the barrier to extract oxygen from the oxide surface via the reaction of CO to CO_2 was too high. The results led them to speculate that it was the interaction of CO with exposed metal that may be responsible for surface decomposition.

The calculation suggested that metal dusting occurs only at the metal cluster sites on the surface. The study explained the pitting phenomena. From this insight, they helped develop a method for monitoring the extent of metal dusting via in-situ FTIR diagnostic tests of the CO stretching vibration on metal. The value in the

calculations here was to indicate what mechanisms were likely and what strategies might be used to monitor and prevent metal dusting.

Adsorbent Structure Identification

One of the first applications of molecular simulation in Air Products was to identify the crystal structure for an important reversible adsorbent powder (RAP) material. Monte Carlo simulation was used with Rietveld refinement to establish the crystal structure of $LiCo(CN)_5 2DMF$ to within 5%. The findings were reported in a past issue of *Chemical and Engineering News* (Feb. 5, 1996).

Si-Al Distribution in Zeolite

Molecular simulation has been used to identify the detailed spatial distribution of Si and Al in Faujasite. The specific zeolite was one that Air Products later patented for the separation for air. ^{29}Si NMR was used in conjunction with simulation to identify the most likely sites of Al. The simulation results demonstrated how Al segregated from Si on different sub-lattices.

DISCUSSION

Balancing Education of Users and Effective Usage

In educating other researchers, one of the premier challenges that the molecular modelers face is managing expectations. The other scientists, engineers, and management can still be somewhat bimodal; i.e., either believe that molecular modeling will not have an impact or that it will solve all the company's problems. For example, many of the bench-top chemists are still reluctant to follow the leads from modeling efforts. The demonstrated usefulness of molecular modeling is slowly leading to more realistic expectations for molecular modeling. Molecular modeling can be used as an effective tool when coupled with the experimental program at the initiation of a project.

The molecular modeling group has helped train over 35 researchers with modeling algorithms, but many of these individuals are still unable to carry out calculations by themselves. Consequently, their momentum slows with time. Much of the work then transfers back to the modelers who have the challenge of prioritizing research needs. While frustrating, it teaches that it is more reasonable to use the specialists for specialized calculations rather than to expect widespread development of such expertise, at least for now.

Needs for Computer Programs

For computer programs to be more useful, improvements in accessibility and portability are needed. Programs are particularly needed for modeling phase equilibrium, but most existing programs are academic codes. Typically, they lack the shells or graphical user interfaces needed to be widely useful, and their source codes may not be open. An ideal for increasing portability would be to break the links between molecular structure, force fields, algorithms, and front-end tools. While available from some companies, there should be more standard agreements for short-term code evaluation. Increasingly, codes must be able to run on a range of modern parallel computing platforms. Finally, the hosts expressed a desire for industrially available, general-purpose Monte Carlo codes.

Technical Needs

Future technical needs were split into computational quantum chemistry and molecular simulation topics (Table B3.2).

Table B3.2
Future Needs in Applied Molecular Modeling at Air Products

Computational Quantum Chemistry	Molecular Simulation
More realistic models for the condensed phase: solvation, solid-state chemistry, heterogeneous catalysis Improved accuracy: energy predictions to less than 2 kcal/mol for systems of the order of 40 atoms *Ab initio* MD simulations: • More efficient schemes • Application to inorganic adsorption systems • MO/MM methods Robust transition-state search algorithms Ten-fold increase in the speed of *ab initio* calculations	Continued convergence of classical/QM methods (a*b initio* MD, QM-derived force fields) Configurational-bias Gibbs Monte Carlo simulation for general PVT and phase behavior: Application to larger and more complex molecules Mesoscale modeling techniques Parallel commodity PCs More accurate, complete force fields: • General procedures to generate FF parameters • Force fields which account for induced polarization • Force fields with charge distribution better than atom-centered sites Computational/combinatorial chemistry methods for automated screening and design

Keys to Success

The successful use of molecular modeling at Air Products has occurred primarily when it was used as an input into technical business decisions, similarly to what has been described by Joe Golab of BP (see BP-Amoco report). This includes the application of fast, low-cost screening methods in order to prioritize experiments. Many times the results of such calculations can be used to indicate whether or not a project has potential (go/no-go decisions). In addition to fast, low-cost screening, molecular modeling has also been used to enhance research activities. Systematic studies are used along with experimental data to develop comprehensive and searchable databases of information for particular systems. The calculations are used to offer insights into controlling chemistry and physics. Molecular modeling provides a useful aid in intellectual property matters by way of substantiating patents or refuting other patents.

One of the keys to success has been the early coupling of experimental and molecular modeling efforts. When modeling is attempted as a last resort where nothing else has worked, chances of success are least, and the only resulting message is that the modeling didn't work. That seems obvious, but people inevitably remember the failure rather than the long odds.

As in any team effort, the expectations from each side must be realistic. Molecular modeling can currently be used fruitfully to provide trends. It probes the atomic structure of a material and therefore offers molecular-level insight. This usually requires some critical level of structural and composition analysis. It is important to respect the knowledge of the clients, who may not initially understand where the modeling methods' strengths and limitations lie. They often need most to know the degree of quantitative accuracy to expect. In turn, the modeler must recognize the connections and priorities among different project objectives. Finally, there must be a good match between the project and modeling time scales. A good calculated result is worthless if it cannot be achieved in time for the project's needs.

Site: **BP p.l.c. (Amoco)**
BP Naperville Complex
150 Warrenville Rd
Naperville, IL 60566-7011
http://www.bp.com/

Interview Date: 12 January 2000

WTEC Author: P.R. Westmoreland (based on the interview, literature, and conference presentations by Amoco and BP personnel)

Interviewee: Dr. Joseph T. Golab, Chemical Modeling Specialist, BP Chemicals

BACKGROUND

BP, based in London, is a multinational oil and chemicals company formed from several British and American companies with long histories:

- In 1909, the Anglo-Persian Oil Company was created for exploration and production of oil in the Middle East, a hoped-for area of petroleum resources. Its founders were principals of the Burmah Castrol Co., which retained an equity interest until 1966. That time span included a name change to Anglo-Iranian Oil Co. in 1935, Iranian nationalization of the oil fields in 1951, negotiated resumption of Iranian production on a 50/50 basis with Iran in 1954, and renaming in that year as British Petroleum. In the 1960s, BP made significant discoveries on holdings in the North Sea and on the North Slope of Alaska.
- In 1987, it acquired SOHIO. This was the former Standard Oil of Ohio, which was the original base of John D. Rockefeller's Standard Oil Trust, broken up by the U.S. Supreme Court in 1911 into regional companies including Standard Oil of New Jersey (later Esso, Enco, Enjay, and the Exxon portion of ExxonMobil), Standard Oil of New York (SOCONY, SOCONY Mobil, Mobil, and then the Mobil part of ExxonMobil), Standard Oil of California (now Chevron), Standard Oil of Indiana (later part of Amoco), Atlantic Refining Company (later ARCO), Continental Oil Company (later CONOCO), Ohio Oil Company (later Marathon Oil), and other components. SOHIO had also been involved with BP through the North Slope production and refining businesses.
- Soon after, in 1988, BP acquired Britoil.
- In 1998, it acquired Amoco, itself formed in 1954 from American Oil Company and Standard Oil of Indiana. The combined company was named BP Amoco during the transition, but it changed officially to BP p.l.c in May 2001. The combined company brought together similar activities in exploration, production, refining, chemicals, and research and engineering development. The chemical businesses were complementary; BP brought leadership in polyethylene, acetic acid, and acrylonitirle, while Amoco Chemicals had great strength in terephthalic acid, para-xylene, and meta-xylene. Overall, the new company had $68.3 billion in revenue in 1998.
- In 2000, BP acquired ARCO, which had become focused on its oil and gas businesses. Also formed from the Standard Oil breakup, ARCO had acquired Sinclair Oil in 1968 and the Anaconda Company (copper production) in 1977, along with other resource companies. In 1997, it had divested its 50% ownership in Lyondell Petrochemicals, sold ARCO Chemicals to Lyondell, and sold most of its coal properties. Its 1998 revenues had been $10.3 billion.

BP is now the third largest oil company in the world. At the end of 2000, its overall revenues were $148 billion, and it had 88,100 employees (America Online MarketGuide 2001).

MOLECULAR MODELING R&D ACTIVITIES

BP had a substantial molecular modeling effort in the UK during the 1980s and early 1990s. The company used molecular modeling to assist in developing catalysts and new products, e.g., brake linings, polymerization catalysts, and corrosion and scale inhibitors, which were all successful applications. This

group and activity were eliminated in the mid-1990s, as discussed in the Royal Society of Chemistry site report.

Keith W. McHenry, Jr. (Alkire 1996) was the champion for the start of molecular modeling at Amoco, aided by Joseph F. Gentile (Manager, Information and Computer Services). Dr. McHenry was a researcher in oil processing, notably in catalysis. In 1974, he became Vice President of Amoco Research and Development, later serving as Senior Vice President for Technology in Amoco Corporation from 1989-1993.

In the late 1980s, a modeling "community of interest" recommended the formation and staffing of a molecular modeling group. As a result, two people were moved into modeling groups in Amoco Oil and three into Amoco Chemicals. At this time only Aileen Alvarado-Swaisgood (now of MSI/Accelrys) was a computational chemist with that specific theoretical background. The others were traditional chemists who had computer skills and/or familiarity with one of several chemistry modeling tools, such as CAChe (Fujitsu).

The Amoco Oil effort, led by Chris Marshall (now of Argonne National Laboratory), was focused on heterogeneous catalysis. Two achievements were analysis of optimum Si/Al ratios for improved cracking-catalyst performance and diagnosis of a catalyst performance problem as being due to pore obstruction by tramp alumina, the latter by interpretation of available 3-D geometry availability within the pores. This group was eliminated and work was discontinued in the early 1990s.

Dr. Golab, who presently leads activities in molecularly based modeling, arrived in 1991 from NCSA at the University of Illinois. His doctorate in chemistry is from Texas A&M University. He became the leader of NCSA's computational chemistry group, which then led to a job at Amoco. He has been active in the modeling community, organizing symposia for the American Chemical Society and the American Institute of Chemical Engineers. In 2000, he was named Membership Chair of AIChE's newly formed Computational Molecular Science and Engineering Forum.

Modelers in Amoco Chemicals and subsequently in BP have covered a wide range of technologies. In the past, molecular modeling focused on molecular structure: properties such as Mulliken population, dipoles, and spectra; thermodynamics; and reaction pathway investigation. At present, there is increased emphasis on kinetics, such as mechanism determination and Fukui factors. Materials modeling also has been oriented to structure and properties with steady interest in crystals, polymer properties using van Krevelen group-based analyses but shifting to a basis in morphology, and recent interest in QSAR methods for high-throughput experimentation. The relative importance of modeling for several areas is listed in Table B3.3.

The incentives for beginning and continuing this type of work have been the need to answer particular questions, proposed efficiencies and cost savings, and the availability of in-house expertise, sometimes accompanied by the need to keep up with competitors and preparation for future needs and capabilities. Decision strategies and several of the Amoco Chemicals' activities have been described in the literature:

- Polyamides based on neopentyldiamine (NPDA, 1,3-diamin-2,2-dimethylpropane) seemed to be attractive polymers. They were transparent and had high glass transition temperatures, but high molecular weight proved difficult to achieve. The literature suggested that cyclocondensation might compete with the desired chain polymerization, becoming a "chain-stopper" that would effectively keep molecular weights down. Analytical chemistry identified the experimental yield of a tetrahydropyrimidine that might play such a role.

  Using the PM3 semi-empirical method, the guiding thermochemistry was calculated for model reactions (Golab 1998), beginning with the reaction of phthalic anhydride and NPDA to make an amic acid amine (AAA). If AAA cyclocondenses to pyrimidine, the chain stops. If AAA's carboxylic OH and amine H dehydrate to give an imide-amine (IA), the chain can continue to grow, although a cyclocondensation side reaction of IA can give another chain-stopping pyrimidine. Gibbs free energies of reaction were calculated. At 300 but not at 475 K, ΔG_{rxn} is negative for the initial reaction forming AAA, indicating that low temperatures are crucial to chain growth. Also, the undesirable cyclocondensation of AAA to pyrimidine proved more negative (more strongly favored thermodynamically) than the parallel reaction of AAA to form IA and, subsequently, polymer.

This combination of data and calculations pointed toward low-temperature polymerization using a more reactive carboxylic-acid functionality than in AAA. Further experiments then identified a carboxylic acid chloride that, at low temperatures, successfully yielded the desired high-MW polyimide (Yokelson and Green 1993).

- Design of an ethylbenzene process was aided by an equilibrium reaction model for alkylation and transalkylation at various temperatures and pressures, constructed with the aid of AM1 semi-empirical calculations (Golab 1998). Gibbs free energies of reaction were calculated involving 12 species, notably the polyethylbenzenes up to tetraethylbenzene. Where possible, these predictions were tested against data, as for the initial alkylation reaction of ethylene with benzene. Free-energy minimization was performed using the reaction-modeling package Chemkin (then available from Sandia, now from Reaction Design, Inc.). Agreement of the predicted and measured ratios of polyethylbenzene to monoethylbenzene was impressive, leading to its acceptance for use in process design.
- Another interesting problem was development of improved moisture and oxygen barrier for food-packaging film (Golab 1999). Passive barriers operate by slow permeability, but the goal was an active barrier that would act as a trap. The Amoco products AMOSORB™ 1000 and 2000 had proved successful for many applications, but the dispersed inorganics in these films make them opaque to translucent in black or gray. This was deemed unacceptable for food packaging, where the consumer wants to be able to see and examine the food through the wrap.

  Again, GAMESS was used to carry out AM1 calculations of the enthalpy and Gibbs free energies of pertinent reactions. A substrate was identified with better trapping kinetics, leading to a clear organic film for food. Patents were applied for and production was begun.
- To help design improved metallocene catalysts for manufacturing elastomeric polypropylene, molecular mechanics and electronic structure calculations were applied (Bormann-Rochette and Golab 1999; Golab 1998) in concert with experiments. Two stereochemical isomers of the catalyst are proposed to correspond to polymers of two different tacticities. A meso isomer produces atactic, amorphous polypropylene segments that can flex between crystalline blocks of isotactic polymer, generated by the rac form of the catalyst. This molecular structure allows elastic behavior of the product polymer.

  The two catalyst isomers were modeled, examining the more stable isomer for various organic ligands attached to the catalytic zirconium. Three molecular mechanics methods were used: Rappe's RFF reaction force field and software, UFF with Cerius^2 (MSI), and ESFF with Insight/Discover (MSI). DFT energies at RFF geometries were calculated, as were HF/3-21G energies and geometries with NWChem (PNNL). RFF seemed to predict interactions best among the force-field methods, compared to data for different ligands, so it was used to correlate tacticity as quantified from ^{13}C NMR. The methods were then used for further experimental development of new catalysts.

More detailed description of thermochemistry calculations for design is provided in Golab and Green (1998).

Molecular modeling at the heritage-Amoco part of BP has been carried out by one full-time staff member (Dr. Golab) and several who use modeling in their work without help or supervision (over time including Mike Green, Craig Bailey, Paul Cahill, John Couves, Alex King, Jim Kaduk, Eric Moore, Andy Ernst, and Eric Ziegel). At any given time, the chemistry modeling "store" has at least four open (active) projects (Golab 1999).

At the time of this interview, Dr. Golab was reluctant to say that molecular modeling was established within BP. Internal successes had made chemistry modeling accepted as a "niche" or "emerging" technology, felt to be successful and valuable. However, it was not yet classified as a technology that is "needed to win."

Table B3.3
Importance of Molecular Modeling for Predictions

Quantity	Level of Importance
Ideal-gas thermochemistry	Extreme
Ion chemistry and equilibrium in solution	Very
Gas kinetics	Very
Heterogeneous catalysis	Very
Heat-transfer properties	Very
Mass-transfer properties	Very
Condensed-phase thermochemistry	Medium
Solution thermochemistry and mixture activity relations	Medium
Surface binding (adsorption)	Medium
Homogeneous catalysis	Medium
Fluid-mechanics properties	Medium
Crystal morphology from solution	Medium
Chemical analytical methods	Medium

MODELING AND HARDWARE AND SOFTWARE

At the time of the interview, hardware available at BP for molecularly based modeling could be classified into three categories:

1. Silicon Graphics Indigo, Power Indigo^2, O2, Power Challenge (parallel), Origin (parallel)
2. Parallel supercomputer cluster of HP-UX 9000/755s (donated to IIT in 2000)
3. Pentium PCs, singly and in a parallel Beowulf cluster

For the expert, it is very important to use both personal computers and stand-alone workstations for modeling with a minimum need for interaction with internal information-technology organizations; for the nonexpert, it is very important to use personal computers. Soon (within 12 months), all calculations will be carried out by parallel computation of various kinds. Single processors will be used for visualization of input or results. Multi-processor computers are easier to maintain and use, but distributed technologies, especially of multi-processor PCs, have appeal in an industrial setting because there are lots of cycles available in off-shift hours.

BP uses a range of modeling codes (Table B3.4), requiring a software budget of about $100,000 per year. Almost none are in-house codes. Molecular Simulations, Inc. (MSI, now Accelrys), a subsidiary of Pharmacopeia, Inc., is BP's major software vendor. BP belongs to the MSI Catalysis & Sorption Consortium (C&SC), in which a major portion of its software development is performed. In addition, the group has several computational chemistry programs at its disposal. Dr. Golab found it ironic that the code with the most use was GAMESS-US, for which BP had little, if any, say in development. Most of BP's software is running on scientific UNIX workstations with superior graphics capability; e.g., Silicon Graphics computers. However, as forefront chemistry modeling software companies begin to offer powerful applications for the PC that rival their UNIX equivalents, BP researchers expect to migrate towards PCs.

Table B3.4
Computer Codes in Use at BP as of June 2001

Application (Platform)	Provider	Brief Synopsis of Uses
Cerius^2 (UNIX)	Molecular Simulation, Inc./Accelrys	General molecular modeling, molecular mechanics, molecular dynamics, quantum mechanics, QSAR, polymers, catalysis
Chem4D DRAW (PC)	ChemInnovation Software	Molecular drawing tool kit for office tools like MSWord
CHEMKIN (UNIX or PC)	Sandia National Laboratories / Reaction Design, Inc.	Chemical equilibrium
GAMESS (Parallel UNIX & LINUX)	Gordon Research Group, Iowa State Univ. and Ames Laboratory	Quantum mechanics, thermodynamics, kinetics, solvation, spectra
Gaussian 98	Gaussian, Inc.	General molecular modeling, quantum mechanics, thermodynamics, kinetics, solvation, spectra
Materials Studio (PC)	MSI/Accelrys	Cerius^2 migration to PC
MCM (UNIX & PC)	Rappe Research Group, Colorado State University	Molecular mechanics, simulated annealing, molecular dynamics
NIST Database 25 (PC)	NIST	Thermodynamic properties database for chemical compounds; predictions via Benson
NIST Database 17 (PC)	NIST	Kinetic reaction database
NWChem (Parallel UNIX)	Pacific Northwest National Laboratory	Quantum mechanics, kinetics, thermodynamics, spectra

Choice of method used is driven by multiple factors. Table B3.5 sheds light on these factors by reviewing the relative usage of different methods. (Note that these projections strongly reflect the particular business interests of BP and the experience with different modeling methodologies at the time of the interview.)

Software development is done by collaboration. Dr. Golab is a Pacific Northwest National Lab (PNL) affiliate scientist and as a result, he enjoys an informal collaboration with the PNL's Theory, Modeling, and Simulation Group. In addition, BP/Amoco and MSI have traditionally worked together toward better software and/or specific applications. These collaborations are both formal (proposal, money, and results are exchanged) and informal.

OPERATIONAL AND STRATEGIC ISSUES

Role of Molecularly Based Modeling, Including Integration

Dr. Golab remarked that it is important to realize that molecularly based modeling is rarely employed single-handedly to solve a problem. The "big impact" problems are really big—their efficient solutions depend on getting the right people together, doing the right things, at the right time, for the right reasons. Exceptional results from individual contributions will become less important to an organization. The real challenge will be the coordination, integration, and deployment of a focused, multidisciplinary team to turn disparate information into organized, exploitable, implementable knowledge.

Table B3.5
Methods in Use at BP and Projected Use

Method	Extent of use now	In the near term (within 2 yrs)	In the long term (within 7 yrs)
Wavefunction-based *ab initio* quantum chemistry	High	High	High
Electronic density-functional theory	Low	Low	Medium
Semi-empirical molecular-orbital theory	Medium	Medium	Low
Molecular mechanics	Low	Low	Medium
Molecular simulations of fluids	Medium	Medium	High
Atomistic simulations of solids	Low	Low	Medium
Spatially or temporally hybrid methods (combination of electronic-structure and/or molecular-simulation methods)	Low	Medium	High
Mesoscale modeling by united-atom or chemical-cluster modeling	Low	Low	Low
Theoretically based correlations	Medium	Medium	High
Empirically based correlations QSAR or QSPR	Medium	High	High
Chemoinformatics, bioinformatics, and data mining	Medium	High	High

Successful impact depends on early involvement in the development process, but that also depends on motivated interest by the rest of the development team, usually driven by past contributions of modeling. Modeling should be part of an RT&D project from "day one." Chemistry modeling helps (a) discover new products and processes, (b) shorten development time, and (c) save money. It has even lead to innovation (implementation of an unobvious idea). It is vital to continue to show that molecular and materials modeling tools are helpful in solving practical chemical problems and can be integrated into legacy ways of solving problems.

Organizationally, it is important that industrial managers possess an adequate scientific background and use it in "value" decisions for the company.

Identifying and Selecting Projects

In the beginning, BP modelers decided what projects to work on in conjunction with the primary client. Then they marketed their skills at internal technology conferences and worked on projects that were funded by the primary client. In many ways, this model is still how they work; however, they no longer have to market their skills—"We have plenty of work." Having advocates among management or clients who recognize the appropriate uses of the methods is imperative to success and future work. The best advertisement for chemistry modeling is a satisfied client.

In this sense, mergers, multiple locations, and multi-national components offer real opportunities. They enhance the project portfolio while communicating effectiveness of the methodology to a wider audience.

Criteria for Success

Dr. Golab notes multiple ways of assessing success at BP:

- First and foremost, did the project preclude involved, preliminary experimentation that traditionally might have been done? Did the model predict or improve the yield, selectivity, or waste reduction for a given chemical process? How well do the molecular-level characteristics of a model compound compare to its macroscopic properties and behavior?
- Second, does or can the model provide valuable, valid, and relevant information to the primary client in a timely manner? Is this a one-time model or can it be used again?

- And finally, did the modeler and primary client collaborate closely? Did a beneficial exchange of information take place during that interaction? Could it have or did it lead to an innovation, the implementation of an unobvious idea?

Big successes (models that are re-useable, results that impact a product line, ideas that are transferable among business units) are more important at this time than incremental achievement. In the BP system, a modeler's time must be charged out to a business unit, presumably doing modeling work. An indicator of success will be when modeling is seen as a crucial part of the company's successes rather than simply being required to maintain a "net zero" cost to the company.

Identifying and Educating Colleagues Who Need to do Modeling

At a basic level, it is important for people to know the language of computational chemistry; i.e., what is Hartree-Fock, what is a heavy atom, etc. Those who want to use the tools effectively or who have come to expect the tools to be available (recent graduates, for example) should be trained in the application of interest. However, in the future, individual contributions will be less important. The coordination, integration, and deployment of a focused team for problem resolution will be the norm.

People decide that they need to do modeling because they see it as having been useful for other researchers. To educate new users, a mix of internal courses and/or external speakers is used.

Interfaces and Roles of Software Companies

The software interface is inevitably important. For the expert or specialist, if there is an interface, it ought not to rely on tools outside itself (like setting up an input deck that needs "a little" tweaking using UNIX editors "jot" or "vi"). The interface should be self-contained for any level of user. Otherwise, there is too much to know and learn besides the pertinent chemistry or engineering.

Expert systems or problem-solving environments would be especially helpful when a problem is outside the specialist's area of expertise. For example, consider an expert in quantum mechanics who must use a molecular mechanics technique. Not only would help be useful for problem setup but also to identify what force field would be best for the problem, how long to expect for convergence, and so on. These same systems would be very helpful for nonexperts.

FUTURE AND CONTINUING NEEDS

Looking to the future, Dr. Golab sees a number of applications for molecularly based modeling that will impact the businesses of BP and its competitors. A general trend will be use of lower-cost feedstocks, aided by more efficient conversion of synthesis gas to hydrocarbons.

Process design and manufacturing is another such area. In the next five years, he foresees the following:

- Integration of molecular and engineering modeling
- Integration of "experimental" and "calculational" data
- The design of efficient chemical processes from microscopic molecular behavior
- Crude process monitoring based on modeled specifications

Meanwhile, longer-term applications include guidance for producing multiple products from similar process(es) within a chemical plant's existing footprint and more sophisticated process monitoring based on modeled specifications.

In catalysis, over the next 5-10 years he expects to see more efficient, longer lived, more specific, more active, less expensive catalysts. This expectation hinges on better understanding or redevelopment of the process because many catalysts operate at their "theoretical" limit now. Another advance should be to link catalyst design to final materials' microstructure.

Dramatic opportunities should arise in material science and technology. Dr. Golab's comments emphasized materials used for their mechanical properties. He breaks these advances into those likely in the next five years:

- Product tailor-made for clients' needs—process directly effects microstructure and/or properties
- Better understanding of relationship between molecular and bulk structure
- Prediction of toxicity and environmental fate from chemical structures
- Better prediction of properties in general (like boiling or melting points)

Advances likely in the next ten years include the following:

- Understanding the difference between predictive properties and actual performance
- Replacing traditional materials (e.g., glass, metal, wood) with polymers
- Using the lowest cost feedstocks for high-tech materials

In the next fifteen years, Dr. Golab foresees "intelligent" materials will be created that carry or deliver information, similarly to the freshness meters on some dry cell batteries today.

To take advantage of these opportunities, Dr. Golab generalized that the highest-priority needs for better modeling within BP were (1) more accurate force fields and interaction potentials and (2) the ability to model transition metals accurately. At a second level of high priority were the ability to do bigger problems, increased accuracy for current-sized problems, and the ability to handle larger amounts of results.

It is possible to project some of the specific modeling technologies (theories, methods, codes, and computer systems) that will be important for such molecular and materials modeling in 5, 10, and 15 years to satisfy these applications. Dr. Golab's assessments are presented in Table B3.6.

At the same time, Dr. Golab contends that conventional molecular and materials modeling technologies cannot meet the requirements of these applications in the next fifteen years. From the most pessimistic perspective, if things do not improve, industry will abandon molecular and materials modeling technologies as a ubiquitous tool and keep it at the specialists' level, if it is kept at all.

There are a number of changes from the present conventions that he expects to make the difference:

- There will be a shift in the interface to be more Web-browser-based.
- There will be a change in the nature of the human interaction with a program using chemically based voice recognition. For example, in the future, a chemist could say, "Using benzene as reagent A, calculate the Hartree-Fock Gibbs free energy of simple alkylation using the list of chemicals in file 1 as reagent B." A program will then generate and present the results. Modeling tools will be more accessible to all scientists with very little "learning curve" dilemmas.
- Experimentation, becoming much more combinatorial, will rely on molecular and materials modeling tools for direction and analysis. Combi-chem is in the very early stages of application. The whole modeling group (statistical, chemical, and engineering) is involved in the enterprise.
- Databases and what they contain will take on a much greater importance.
- Processes will also rely on models and "instant comparison" to theoretical limits to ensure product specifications.

From BP's perspective, the best way to facilitate creation and transfer of new science to technology is to drive it with "I want to..." scenarios. By defining strategic problems based on "I want to" scenarios that are important to industry and/or government, modeling technology development and transfer can be identified and funded effectively. This defining stage must include scientists and engineers, manufacturers, and distribution folks. An "I want to" is a very high, pie-in-the-sky, worry—similar to the Grand Challenges, only more practical. An example "I want to" would be, "I want to create 9 tons of gebiff at 95+% yield within a fortnight in the year 2004."

Table B3.6
Near and Longer-Term Modeling Technologies Expected to Have Impact on Applications

Timeframe	What	Why
5 years	HF and DFT extended to molecules of 1000 atoms in large basis sets	"Simple" thermodynamics and property calculations crystal structure manipulation
	MCSCF and CC extended to molecules of 100 atoms in large basis sets	Accurate calculations of thermodynamics and kinetics properties
	CI extended to large molecules (50 heavy atoms)	Kinetics
	Internal parallel and/or clusters of PCs	Cost-effectiveness
10 years	Routine HF and/or DFT for any molecule (sunset of semiempirical methods conjecture)	Thermodynamics and property calculations, QSAR for HTE
	MCSCF, CC, and/or CI extended to molecules of 1500 heavy atoms in large basis sets	Calculations of thermodynamics and kinetics properties with accuracy near or at the same level as experiments
10-15 years	Problem Solving Environment	"Bench" chemist will be able to think about the chemical problem without as much know-how needed to formulate input
Ongoing	Parallel networking of PCs	Use of corporate PCs during down time (at night)
	Much better force fields for MM-based methods	"Simple" sketching out of problem
	Much better force fields for QM/MM methods	Real problem solving (reaction on surface)

In summary, Dr. Golab made the following remarks:

> Assuming that molecular and materials modeling tools continue to improve over time (as they have), the future of the technology is very bright. For the most aggressive companies, modeling will provide a tangible competitive advantage. This is mainly because, if implemented correctly, modeling involves comparatively little human time and effort relative to amount of information that can be obtained versus other means.
>
> It is further anticipated that the primary client will perform initial experimental and computational developmental work. As a result, future experiments will rely on modeling results to target specific product development and/or improvement. It is possible that a majority of "new" experiments will be done to improve models. Future modeling projects will be much more sophisticated in scope, resource demands, and subsequent bottom-line achievement. As more and more experimentalists use modeling to support and justify their work, "Modeling" and "Experiment" labels on data will become meaningless.
>
> The biggest impact will come when modeling data is considered to be just as reliable as experimental data. At that point, models will be used to explore the limits of what is possible—beyond what is currently doable (e.g., extreme conditions) reliably. In the future, chemical process models (routes to materials) will be fully integrated across disciplines, fully detailed, and fully reliable.

CONCLUSIONS

The history of molecular modeling activity in BP and its historical components has mirrored the field's problems with over-optimism and the subsequent over-compensation for not living up to ambitious visions. Such work in the heritage-BP and heritage-Amoco Oil components was once vigorous and scored real successes. Nevertheless, both groups were eliminated. The heritage-Amoco Chemicals component has survived by keeping closely tied to the whole development process, as well as by using low-level methods when low-level methods are adequate and high-level methods where those are needed. With the continued

growth of BP, this strategy and its successes promise to engender more success, increasing the use of modeling within the company.

REFERENCES

Alkire, R.C. 1996. Keith W. McHenry, 1928-1994. *Memorial Tributes: National Academy of Engineering*, Vol. 8, pp. 148-152. (http://books.nap.edu/books/03090555X/html/148.html).

America Online MarketGuide. 2001. http://aol.marketguide.com/mgi/.

Bormann-Rochette, L., and J.T. Golab. 1999. National Meeting of the American Chemical Society, Anaheim CA, March 22, 1999.

Golab, J.T. 1998. Making industrial decisions with computational chemistry. *Chemtech*, 28(4), 17.

Golab, J.T. 1999. Computational molecular Science at BP. National Meeting of the American Chemical Society, Anaheim CA, March 22, 1999.

Golab, J.T., and M.R. Green. 1998. Practical chemistry modeling applied to process design studies. In *Computational Thermochemistry: Prediction and Estimation of Molecular Thermodynamics*, edited by K. Irikura and D. Frujip (ACS Symposium Series 677, Washington, DC).

Yokelson, H.B., and M.R. Green. 1993. *Polym. Mater. Sci. Eng.*, 69, 84.

Site: **CambridgeSoft.com**
100 CambridgePark Drive
Cambridge, MA 02140
800 315-7300 / 617 588-9300, Fax: 617 588-9390
http://www.cambridgesoft.com/

Dates of Interviews: 17 February and 9 March 2000

Interviewer: P.R. Westmoreland (report author)

Host: Bruce R. Gelin, PhD, Business Partner Manager, bgelin@camsoft.com

BACKGROUND

CambridgeSoft.com was founded in 1986 as Cambridge Scientific Computing by Stewart D. Rubenstein and Michael J. Rubenstein, changing its name to CambridgeSoft Corporation in 1996. Originally, it focused on developing ChemDraw, a Macintosh-based program for drawing two-dimensional chemical structures. Not only could a user generate publication-quality images using conventional organic-chemistry structural representations, but the structural descriptions used to build the image could be used in computer databases. Later versions added formula and elemental composition calculation, IUPAC naming, and correlation-based estimation of proton and ^{13}C NMR shifts, boiling point, melting point, critical temperature, critical pressure, Gibbs free energy, logP, refractive index, and heat of formation.

Subsequently, the software line was expanded to include molecular modeling and database management. Chem3D, the molecular modeling code, includes MM2-based energy minimization and molecular dynamics, MOPAC semi-empirical molecular orbital calculations, and an interface to the Gaussian 98 *ab initio* code. The ChemFinder database manager allows searching by structure, substructure, similarity, text field, or numeric field and is a front end for preparing queries to other databases. The combination of products (ChemOffice) is intended for routine use by chemists to record and communicate their work.

Hardware applicability was expanded to include Windows-based computers. The company finds Windows use to be prevalent in U.S. and European companies, while Mac is strong in education and especially in Japan.

The company's business strategy currently has a strong Internet focus, including sales portals. Its marketing and technical publication, ChemNews.com (http://www.chemnews.com/), is available in print and on the Internet in English, French, German, and Japanese. It hosts http://www.chemfinder.com/, a free portal to chemical properties and extensive information on individual chemicals. The existing desktop tools are being adapted to work in Internet settings, developing the database portion into server software and producing versions of ChemDraw and Chem3D that act as the network clients. ChemDraw, for instance, is available as a free browser plugin and as a Java application. These developments should facilitate chemically intelligent Internet applications.

The company estimates it has about 250,000 paid users, including scientists from Merck, Bristol-Myers Squibb, Procter & Gamble, Abbott Laboratories, Agouron Pharmaceuticals, American Home Products (Wyeth-Ayerst, Lederle), Glaxo Wellcome PLC, Johnson & Johnson (Ortho McNeil) , Mitsubishi, Pfizer, Vertex Pharmaceuticals, and Pharmacopeia.

DISCUSSION

Dr. Gelin is a PhD chemist who studied at Harvard with Martin Karplus. Prior to joining CambridgeSoft, he had worked at Polygen. That company was founded in 1984 to develop and sell the QUANTA/CHARMm. It was merged in 1991 into Molecular Simulations, Inc.

For a company with an Internet strategy, issues include bandwidth and coping with new server software (Windows NT and Microsoft's IIS) and Web development tools. Internet applications that run beautifully over local ethernets or fast Internet pathways are almost unusable when the means of communication is a 56K modem. Likewise, adapting codes to new standards for operating systems, Java, and Web markup languages requires a lot of work simply to maintain capabilities.

CambridgeSoft appears to be positioning itself as a chemical information supplier as well as a provider of research utilities. Company representatives note the problem of making non-computerized information available and usable on computers. There is a vast amount of information stored on paper in older journals, company notebooks, manuals, and catalogs that is quite hard to get into computers. In some cases, the only way to extract the chemical content is to key in the information; merely scanning page images facilitates electronic distribution but doesn't capture the chemical significance.

They perceive the future for computer use in chemical education as bright: "College students are increasingly exposed to chemistry software and databases, and soon, students won't remember a world without computers and the Internet!" This does not presume that all chemists will be directly using computational chemistry, but they believe that chemical utility software such as ChemDraw and ChemOffice will be used routinely.

Dr. Gelin described the company's point of view by remarking that "The Internet, with its ability to distribute and collect information, is simply astonishing for its potential and its rate of development. It is perfectly in tune with the increasing globalization of research enterprises. Enabling worldwide collaboration and worldwide access to databases is already important, and it shows every sign of becoming more important to researchers."

Future issues include remote and outsourced use of the software. CambridgeSoft's software has been present only on desktops, often ordered directly by the scientists. With the Internet developments just mentioned, the company is moving into group and enterprise problem-solving, although it is premature to make any assessment of that activity.

Site: **Chevron Corporation**
Petroleum Energy and Environmental Research (PEER) Center
20970 Currier Rd
Walnut, CA 91789
http://www.chevron.com/

Date Visited: 8 December 1999

Reporter: P.R. Westmoreland

Contact: Yongchun Tang, Director of Petroleum Energy and Environmental Research Center at California Institute of Technology, tang@peer.caltech.edu

BACKGROUND

Chevron Corporation is an oil and gas company with sales of $51 billion and net income of $5.2 billion in 2000. It is also half-owner of Chevron Phillips Chemical Company, which is described in a separate report ("Phillips Petroleum Company/Chevron Phillips Chemical Company"). It is due to complete a merger with Texaco in June 2001 to form Chevron Texaco Corporation, a company with $78 billion in assets.

MODELING R&D ACTIVITIES

Dr. Yongchun Tang has conducted molecularly based modeling for Chevron since the early 1990s. Presently he is Director of Petroleum Energy and Environmental Research Center at the California Institute of Technology. He is also involved in some DOE and NSF projects and was to teach at Caltech in spring 2001. He also has an adjunct professorship at Cornell University.

Work began in 1990, focusing on catalysis and oil field chemicals. There have been two people doing this modeling, mostly using software from MSI (Cerius^2) and Schrödinger (Jaguar). Achievements include developing models for the following:

- Predicting the performance of different corrosion inhibitors (Ramachandran et al. 1996)
- Predicting the efficiency of different scale inhibitors using QSAR correlation (Blanco et al. 1997)
- Predicting the performance of different wear inhibitors, including shear stress, viscosity, and other properties (Jiang et al. 1997)
- Predicting the rate of product formation for gasoline reforming kinetics
- Geochemical isotope fractionation (Tang et al. 2000)

DISCUSSION

Chevron researchers have found the biggest technical challenge to be lack of experimental data useful for modeling. Leverage is the key issue. At the same time, a quasi-technical problem is that modeling is not a mature technical area. Finally, in terms of strategy, it is critical for them to leverage their results.

Their biggest needs match several current developments well. Greater computational speed and abilities to perform mesoscale modeling would help them greatly. Likewise, they are very enthusiastic about recent developments in quantum mechanics and molecular simulations of very large systems, including calculations for the presence of solvent.

REFERENCES

Ramachandran, S., B.L. Tsai, M. Blanco, H. Chen, Y. Tang, and W.A. Goddard III. 1996. Self-assembled monolayer mechanism for corrosion inhibition of iron by imidazolines. *Langmuir* 121: 6419.

Blanco, M., Y. Tang, P. Shuler, and W. A. Goddard III. 1997. Activated complex theory of barite scale control processes. *J. Mol. Engrg.* 7, 491.

Jiang, S., R. Frazier, E.S. Yamaguchi, M. Blanco, S. Dasgupta, Y. Zhou, T. Cagin, Y. Tang, and W.A. Goddard III. 1997. The SAM model for wear inhibitor performance of dithiophosphates on iron oxide. *J. Phys. Chem.* B 101: 7702.

Tang, Y., J.K. Perry, P.D. Jenden and M. Schoell. 2000. Mathematical modeling of stable carbon isotope ratios in natural gases. *Geochimica et Cosmochimica Acta* 64(15), 2673-2687.

Site: **The Dow Chemical Company**
Central and New Business Research and Development
1776 Building
Midland, MI 48674
http://www.dow.com/

Date Visited: 7 December 1999

WTEC Attendees: P.T. Cummings (report author), R. Miranda, E.B. Stechel, P.R. Westmoreland

Hosts: Tyler B. Thompson, Research Partnership Leader, tbthompson@dow.com
Dawn L. Shiang, Resource Leader, Computing, Modeling & Information Sciences, dlshiang@dow.com
Nelson G. Rondan, Leader, CMIS, ngrondan@dow.com
Michael H. Mazor, Research Leader, CMIS
Sudhakar Balijepalli, Senior Research Engineer, CMIS, sbalijepalli@dow.com
Mark H. McAdon, CMIS, mcadon@dow.com
Jonathan D. Moore, Research Engineer, CMIS
Joey Storer, CMIS (by conference call from Minneapolis)
David J. Frurip, Research Associate, Reactive Chemicals / Thermal Analysis / Physical Properties, DFrurip@dow.com
Joseph R. Downey, Jr., Research Associate, RC/TA/PP JRDowney@dow.com

BACKGROUND

Dow Chemical Company is a global science and technology company that manufactures chemical, plastic and agricultural products in 168 countries around the world. Dow conducts its operations through 15 global businesses employing 39,000 people. The company has 123 manufacturing sites in 32 countries and supplies more than 3,500 products. Annual sales in 2000 exceeded $23 billion. On February 6, 2001, Dow Chemical Company announced the completion of its merger with Union Carbide Corporation.

Dow spends just under $1 billion/year in research and development. At the time of this WTEC visit, it had four major research and development sites: Midland, Michigan (33 research and development facilities supporting specialty chemicals, plastics, and agricultural products), Freeport, Texas (home of Texas Research and Development as well as several technical service and development laboratories with more than 1000 researchers, chemists, scientists, engineers and technologists supporting such businesses as plastics, epoxies, hydrocarbons, and chlor/alkali), La Porte, Texas (in support of manufacturing activities in polyurethanes, fabricated products, and engineering plastics) and San Diego, California (industrial biotechnology to produce chemicals and proteins from renewable resources). As of early 2001, research personnel had been consolidated at the Midland and Freeport locations.

MOLECULAR MODELING R&D ACTIVITIES

Molecular modeling began at Dow in the early 1980s with the hiring of four people (Dr. Rondan, Dr. Pat Andreozzi, and two people who are now gone). Dr. Jozef Bicerano joined in 1986 and Dr. McAdon in 1987. The early work included studying photodegradation of the fire-resistant polymer polybenzoxazole, indicated to proceed by singlet O_2 attack on the aromatic ring.

Currently, most molecular modeling personnel are located in Midland, but consulting with operating units takes the molecular modelers throughout the company. For example, on the day of the WTEC visit, two members of the molecular modeling team, Jozef Bicerano and Chris Christenson, were both visiting manufacturing sites. Such direct contact is considered essential.

From the beginning of molecular modeling at Dow, there has been strong encouragement to be predictive rather than descriptive. Several examples were given. One example concerned a request by a customer for Dow to produce an organic compound which, along with other requirements, was very stable at 460°C and had a low dielectric constant (<2.7). Molecular modeling at the structural level (i.e., not using QSAR), carried out by Nelson and Rondan, resulted in several candidate compounds. The dielectric constant was predicted using a theory and code developed in-house by Joe Bicerano. For one particular candidate material, the code predicted a dielectric constant of 2.9. It had been measured experimentally as 2.3. Re-measurement showed that the predicted value was correct. The time from request by the customer (June) to delivery of the product to the customer (December) was six months. This short time frame was made possible through the use of molecular modeling.

A major part of the success of molecular modeling at Dow has been in the areas of homogeneous and heterogeneous catalysis. Some of these activities have been reported in patents (LaPointe et al. n.d.; McAdon et al. n.d.; Nickias et al. n.d.) and in the open literature (Feng et al. 1999; Ruiz, McAdon, and Garcés 1997). In this area, molecular modelers work closely with experimentalists to modify catalytic materials in order to optimize performance.

Support and interpretation of analytical measurements has proven valuable. MSI software has been used in conjunction with EXAFS, HRTEM, IR spectroscopy, Raman spectroscopy, NMR shifts, and X-ray diffraction.

Dow has begun a concerted effort in bioinfomatics in support of its agrosciences division. Dow is implementing an Oracle-based system from Pangea Systems (now called DoubleTwist, Inc.) and the Wisconsin Sequence Analysis Package from Genetics Computer Group (now an MSI company).

Dow is a leader in the use of computational quantum chemistry for thermochemical calculations. Dave Frurip, a member of the laboratory involved in producing the JANAF tables of thermochemical properties (Dow 1971), expressed the opinion that computational chemistry for the calculation of change in enthalpy of reaction, a crucial property in safety analysis, has reached "maintenance mode." His group has developed confidence that they can obtain accurate values from computational chemistry calculations. There is a significant cost advantage to the computational chemistry approach. In 1996, it was estimated that experimental measurement of one heat of reaction cost approximately $70,000 compared to $20,000 for a G2 calculation on the same molecule. Now the comparison has moved to $100,000-plus for the experiment versus $2,000 for a G3 calculation. This leverage is great, but it would be even better if a structure could be drawn, desired precision stated, and then the program would go and do whatever calculation was best.

Joe Downey described Dow's current efforts and future expectations for obtaining physical properties for process design applications, both *de novo* and adaptive design—heat capacities, densities, viscosity, thermal conductivity and diffusivity, all primarily in condensed phases, phase equilibria, and critical points to within 3K in the critical temperature and a few bar for the critical pressure. Molecular simulation is beginning to provide a route to many of these properties.

Computing power was recently expanded at Dow by implementing a new, sustainable, high-performance computing (HPC) strategy. This strategy required increasing internal technical expertise for computer support and for special projects. Several staged HPC implementations, including UNIX and Linux, were carried out at two sites (Midland and Freeport) in order to support R&D computing. Current emphasis is on automation, including automated job queuing. These implementations allowed the phase-out of older, more expensive equipment. Hardware and maintenance costs have been substantially reduced.

The software used in molecular modeling at Dow includes Gaussian, Wavefunction's Spartan, MSI's Insight/Discover and Cerius^2, and Schrödinger's Jaguar codes. Dow is a member of MSI's polymer and catalyst consortia. It is moving increasingly to academic codes and the codes produced as a result of government-academia-industry consortia, such as the multi-scale modeling code MPFLOW developed under the BRITE-EURAM partnership and the ESPRIT codes of Theodorou, Ottinger, and Keunings. Dow relies significantly on interaction with academia.

DISCUSSION

It was evident to the WTEC attendees that the molecular modeling activity at Dow was large, mature, successful and growing. In addition to presentations and discussions from the Dow attendees showcasing successes in molecular modeling, Tyler Thompson gave other evidence of the importance Dow Chemical attaches to molecular modeling. Dr. Thompson pointed out that Dow had committed significant resources to participate in government workshops related to molecular modeling and had promoted to high levels in the company successful molecular-modeling practitioners. Molecular modeling enjoyed strong support at the vice presidential level of the company and is in the process of moving beyond the need for a champion by becoming ubiquitous in Dow research and development projects. Recognition of the value of molecular modeling was reflected in molecular modelers being included on patents and on two occasions being included in Dow Inventor of the Year awards.

However, much of the activity at Dow is not reflected in external publications or presentations. Publication is encouraged; the view of the modelers is that they publish results in the open literature when it provides a valuable service to the broader community. As with most companies, results cannot be published that would negatively impact Dow's competitive edge in product development. Other issues can affect publication by companies as well. A good example is when a journal reviewer from the *Journal of Organic Chemistry* insisted on more experiments to support a model. Performing more experiments was not justifiable for the company, so the submission was withdrawn and published instead in *THEOCHEM.*

Several lessons have been learned at Dow Chemical in making molecular modeling successful. The ability to be flexible was a key—the statement was made that there are "no successful purists in industrial molecular modeling."

General communication of modeling results is another aspect. For example, there is an internal web site that documents molecular modeling successes and includes VRML images of molecules.

Another aspect is digesting the results so as to be useful and persuasive. In a recent case where patent claims on adhesives were extended by modeling, Dr. Shiang explained that the results were digested into an Excel spreadsheet. Such a step aided the translation of numbers into explanations. As Dr. Thompson noted, to build comfort with modeling, client researchers need to test models and find that they capture believable insights.

Give-and-take with experimentalists is important as well. Dr. Mazor recalled a meeting of scientists working on new ligands for homogeneous catalysts. A workstation was brought into the meeting for a demonstration using Spartan software and 3D glasses. A vigorous discussion ensued as the participants tried out proposed ligands. Indications were very positive, and several patent filings resulted from the session.

CONCLUSIONS

The molecular modeling activity at Dow was clearly one of the largest among the non-pharmaceutical sites visited. It is a mature effort that has a substantial history of important and impressive successes in the company. The diversity of the effort was notable, ranging from computational chemistry through molecular simulation to mesoscale and macroscale modeling. We understood from our discussions that the molecular modeling effort in Dow is growing substantially.

It seemed clear from the discussions that the modelers at Dow felt that the company-sponsored technical advisory board had played an important role in keeping Dow apprised of new developments in molecular modeling and had been responsible for a focus on multi-scale modeling. The high level of structured external consulting activity at Dow was unusual among the non-pharmaceutical sites visited by WTEC.

REFERENCES

LaPointe, R.E., J.C. Stevens, P.N. Nickias, and M.H. McAdon. February 24, 1985. U.S. Patent Number 5721185: Homogeneous olefin polymerization catalyst by abstraction with Lewis acids.

McAdon, M.H., J.T. Patton, P.N. Nickias, R.B. Shankar, F.J. Timmers, B.W. Kolthammer, D.D. VanderLende, and S.M. Ueligger. October 12, 1999. U.S. Patent Number 5965756: Fused ring substituted indenyl metal complexes and polymerization process.

Nickias, P.N., M.H. McAdon, J.T. Patton, B.P. Friedrichsen, J. Soto, J.C. Stevens, and D.D. VanderLende. January 18, 2000. U.S. Patent Number 6015868: Substituted indenyl containing metal complexes and olefin polymerization process.

Feng, S., J. Klosin, W.J. Kruper, Jr., M.H. McAdon, D.R. Neithamer, P.N. Nickias, J.T. Patton, D.R. Wilson, K.A. Abboud, and C.L. Stern. 1999. Synthesis and characterization of cyclohexadienyl-based constrained geometry complexes. *Organometallics* 18, 1159-1167.

Ruiz, J.M., M.H. McAdon, and J.M. Garcés. 1997. Aluminum complexes as models for bronsted acid sites in zeolites: Structure and energetics of $[Al(OH)_4]^-$, $[Al(H_2O)_6]^{3+}$, and intermediate monomeric species $[Al(OH)_x(H_2O)_{n-x}{\cdot}mH_2O]_{3-x}$ obtained by hydrolysis. *Journal of Physical Chemistry B* 10, 1733-1744.

Dow Chemical Company. 1971. Thermal Research Laboratory, D.R. Stull and H. Prophet, project directors, *JANAF Thermochemical Tables*. Washington, U.S. National Bureau of Standards.

Frurip, D.J., N.G. Rondan, and J.W. Storer. Implementation and application of computational thermochemistry to industrial process design at the Dow Chemical Company. Chapter 17 of *Computational Thermochemistry: Prediction and Estimation of Molecular Thermodynamics*, Irikura, K.K., and D.J. Frurip, editors. 1998. American Chemical Society, ACS Symposium Series No, 677, Washington, DC.

Ji, J., S. Narayan-Sarathy, R.H. Neilson, J.D. Oxley, D.A. Babb, N.G. Rondan, and D.W. Smith, Jr. 1998. [p-((Trifluorovinyl)oxy)phenyl]lithium: Formation, synthetic utility, and theoretical support for a versatile new reagent in fluoropolymer chemistry. *Organometallics* 17, 783-785.

Site: **Dow-Corning Corporation**
2200 W. Salzburg Road
Midland, MI 48686-0994
Tel: (517) 496-4000
http://www.dowcorning.com/

Date Visited: 7 December 1999

WTEC Attendees: R. Miranda (report author), P.T. Cummings, E.B. Stechel, P.R. Westmoreland

Hosts: Dr. John (Pat) Cannady, Assoc. Res. Scientist, Global Leader, Intermediates Expertise Center, tel. 517-496-6471, fax 517-496-6243, pat.cannady@dowcorning.com
Dr. Peter C. Qian, Sr. Res. Specialist, Interface Expertise Center, tel. 517-496-6938, peter.qian@dowcorning.com
Dr. Stelian Grigoras (in absentia, but providing written input)

BACKGROUND

Dow-Corning Corporation, headquartered in Midland, Michigan, is an 8,500-employee worldwide company, with 28 manufacturing sites and more than 10,000 products based on silicon chemistry and silicone technology. It was established in 1943 as a joint venture between Corning Glass Works (now Corning, Incorporated) and The Dow Chemical Company, which continue to own equal shares today. It has research facilities in seven countries and hold more than 2200 patents in the United States and 6800 around the world. Its 1998 sales—60% international—were $2.6 billion, its operating income $298 million, and its R&D investment $199 million.

Its extensive product portfolio supplies a variety of industries: automotive, aviation, building construction, chemicals, cleaning products, coatings and plastics, electrical and electronics, food, health and personal care, information management, paper manufacturing, textiles and mold making. Typical products are lubricants, sealing and high-insulating materials, adhesives, cleaning fluids, temperature and UV-resistant coatings, hyperpure crystalline silicon, interlayer dielectrics, and cosmetics.

The company was reorganized in January of 1999 into five business units closely aligned with market sectors, for example construction materials, performance enhancement materials, and silane intermediates.

Much of the R&D function is distributed among a set of expertise centers, each focusing on a particular set of phenomena, materials or process technologies. They are geographically dispersed, located at Midland, Europe and Japan. Each group involves 5-15 investigators. The expertise centers handle diverse projects, two-thirds of which last from 6 months to 2 years, and a third less than 6 months.

MOLECULAR MODELING R&D ACTIVITIES

Molecular modeling at Dow-Corning was started in 1984 with the hiring of Stelian Grigoras (Physics, U. Illinois), who developed molecular mechanics force field parameters and quantum mechanical basis sets for polysiloxanes (Grigoras and Lane 1987; Chou et al. 1987; Grigoras and Lane 1988). Currently he uses quantum mechanics and molecular mechanics methods, in addition to a newly developed method (PALMYRA from Matsim, Inc.) for mesoscale calculations on nanoparticles. By 1994 a group of about 25 persons was routinely using diverse computational techniques, including molecular mechanics and increasingly the quantum mechanical programs of MSI/Biosym and Gaussian. Peter Qian (chemistry, Clemson), joined in 1989 and did physical property and performance prediction by atomistic simulation (MM, MD, MC). He has applied MSI/Biosym (primarily Cerius^2, which was introduced in 1994) to silicon-modified organics and polymer physics. A current task is the calculation of polymer mechanical, thermal, and transport properties. Pat Cannady (chemistry, Univ. of Illinois) is focused on small molecule and intermediates quantum mechanics. About half of the applications are developed within *ab initio*, DFT, and

semi-empirical MO theory; and a third within MM. Solid-state meso-scale modeling, QSPR, and data mining are practiced less but currently are being expanded.

Dow-Corning has developed expertise in several areas: polymer structure and conformation (Grigoras 1987; Grigoras 1990; Grigoras 1991; Grigoras 1992; DeSmedt et al. 1996), optical properties (Wong et al. 1991; Ni et al. 1991), adhesion and interfacial interactions of siloxane networks (Pfister, Woodward, and Grigoras 1992), miscibility properties, environmental and cosmetics applications (Grigoras, Lane, and LeVier 1987; Mazzoni, Roy, and Grigoras 1997; DeSmedt et al. 1996), permeation of gases in polysiloxanes, chemical reactivity and catalysis, AI-aided process design and control, and CFD.

The centers in Midland have 2 full-time and several part-time molecular modeling specialists, as well as 4-10 engineering experts in CFD and process simulation and control, and 3-4 statisticians. They operate an 8-processor SGI Origin 2000, a 4-R8000-processor Power Challenge, three R10000 Indigo 2, two double and single-processor Octane, two O2 dedicated to molecular modeling and ten O2 for CFD and process simulation. The main molecular modeling programs used are Cerius^2, InsightII, Gaussian, and Spartan, which are installed in about thirteen of the computers mentioned, with an annual maintenance cost of over $70,000. The Japanese centers utilize ChemDraw and MoleStudio. The latter is a PC-SGI interface to allow remote usage of the Midland software.

OPERATIONAL AND STRATEGIC ISSUES

The computational group's activities are justified by supporting the company's strategic directions. Modelers are intimately involved with the experimentalist needs within the expertise centers. Method development is lightly pursued, mainly in collaboration with academia both in the US and abroad. Software consortia are not favored, except with academia (Center for Silicon Research, U. Wisconsin-Madison) or government (F.D. consortium at OIT). Those collaborations lead to publications, which are encouraged at the expertise centers.

FUTURE AND CONTINUING NEEDS

Continuing needs are foreseen in the following example applications:

- Predictive thermochemistry, with an increasing role of data mining and applications to process design.
- Mechanical properties of polymer networks such as silicone rubber. Long-time scale dynamics are necessary.
- Interfacial properties, particularly bonding through surface groups. Emulsions, surfactants, micelles, blends, resins, need more attention.
- Prediction of optical and electronic absorption spectra.
- Reaction pathways, kinetics and catalysis.
- Understand and develop approaches to help the progress of nanotechnology.

Some of those needs could be satisfied with more extensive transfer and collaboration with academic groups. However, there are still critical issues to be resolved:

- A general need is for improved linear scaling methods of strongly electron-correlated systems, so that meso-scale modeling of large systems including transition metals becomes routine. The transition from atomistic to bulk behavior is currently not predictable.
- Method validation with appropriately designed experiments is rare. A critical analysis of current results is lacking. Reference standard molecules should include silanes.
- In terms of computation, parallel processing is still undeveloped; commercial software capabilities are too limiting.
- Organizationally, measures of success in molecular modeling are undefined.

CONCLUSIONS

Dow-Corning Corporation has a very successful 15-year history of application of modeling and simulation for their core business, namely polysiloxane chemistry. Recent company reorganization and refocusing of competencies are imposing stronger demands on the computational activities, particularly molecular modeling and simulation. The prediction capability and performance of current theoretical and computational methods must be significantly increased. The prediction of property and dynamic behavior of extended molecular and interfacial systems will rely more extensively on QSPR methods to accommodate the fast turn-around demands of the production environment. The inclusion of molecular modeling in process design and product development will accelerate its acceptance in the company's profit cycle.

REFERENCES

Chou, S.-H., A.J. Freeman, S. Grigoras, T.M. Gentle, B. Delley and E. Wimmer. 1987. Bonding of noble metals to semiconductor surfaces: First-principle calculations of Cu and Ag on Si (111). *JACS* 109, 1880-1.

DeSmedt, A., S. Grigoras, et al. 1996. Predicting the stability of water in oil emulsions: A new mathematical tool. In *Perfum. Und Kosmetik*, 77, 5541-4.

Grigoras, S. 1987. The balance between electrostatic and van der Waals energies and their effect on substituted polysiloxanes. *Polym. Prepr.*, 28, 2, 405-7.

Grigoras, S. 1990. Substituted polysiloxane polymers: Conformation of the pendant groups. *Polym. Prepr.*, 31, 1, 697-8.

Grigoras, S. 1991. Direct computation of statistical weights from conformational analysis for RIS model: Polydimethylsiloxane. *Polym. Prepr.* 32 (3) 720.

Grigoras, S. 1992. Polydimethylsiloxane: Conformational analysis and configurational properties. In *Molec. Model. of Polym.*, (Ed. J. Bicerano, Marcel Dekker, NY) 161-189.

Grigoras, S., and T.H. Lane. 1987. Ab Initio calculations on the effect of polarization function on disiloxane. *J. Comp. Chemistry* 8, 1, 84-93.

Grigoras, S., and T.H. Lane. 1988. Molecular mechanics parameters for organosilicon compounds from ab initio computations. *J. Comp. Chemistry* 9, 1, 25-39.

Grigoras, S., T.H. Lane, and RR. LeVier. 1987. Bioactive organosilicon compounds. A theoretical study of the hormonal activity of 2,6-cis-diphenylhexamethylcyclotetrasiloxane. *Main Gr. Met. Chem.*, 10 (3) 199-205.

Grigoras, W., and T.H. Lane. 1988. Conformational analysis and interpretation of physical properties of substituted polysiloxane and polysilane polymers. *Polymeric Mat. Sci. and Eng.* 58, 493-7. Also in: Silicon-based polymer sci.: A comprehensive resource, (Ed. J. Ziegler, G. Fearon, ACS Washington, DC), *Adv. Chem.* 224, 125-144 (1989).

Mazzoni, S., S. Roy, and S. Grigoras. 1997. Update on eco-relevant properties of selected organosilicon materials. In *The Handbook of Environ. Chem.*, Vol. 3, Part H, (Ed. G. Chandra, Springer-Verlag, Heidelberg).

Ni, Q.-X., S. Grigoras, et al. 1991. Anomalous optical and ESR properties of doped polydiethynylsilane. *Phys. Rev. B—Rapid Comm.*, 44 (11) 5939.

Pfister, W.R., J.T. Woodward, and S. Grigoras. 1992. Silicone adhesive for transdermal drug delivery. *Chem. Britain* 27(1) 43 (1991); also in Pharmac. Technol. 16(1) 42-58.

Wong, K.S., S. Grigoras, et al. 1991. Femtosecond dynamics of the nonlinear optical response in polydiethynylsilane. *Appl. Phys. Lett.* 58 (16) 1965.

Site: **DuPont Pharmaceuticals (now part of Bristol-Myers Squibb)**
E353/134, PO Box 80353
Wilmington, DE 19880-0353
http://www.dupontpharma.com/
http://www.bms.com/

Date Visited: 19 January 2000

WTEC Attendees: P.A. Kollman and P.R. Westmoreland (report authors), P.T. Cummings, K. Morokuma, R. Miranda, P. Vashishta

Hosts:
Carl Decicco, Manager, Chemical and Physical Sciences, carl.p.decicco@dupontpharma.com
Steve Brenner (Contact), Executive Director, Physical Sciences, (302) 695-7354, (302) 695-9090 FAX, Stephen.L.Brenner@dupontpharma.com
Dennis Underwood, Group Leader, Discovery Informatics and Molecular Design, dennis.j.underwood@dupontpharma.com
Pat Brady, Discovery Informatics and Molecular Design, G.Patrick.Brady@dupontpharma.com
Andy Tebben, Discovery Informatics and Molecular Design, andrew.j.tebben@dupontpharma.com
Zelda Wasserman, Discovery Informatics and Molecular Design, zelda.r.wasserman@dupontpharma.com
Chong-Hwan Chang, Protein Crystallography, chong-hwan.chang@dupontpharma.com
Paul Meenan, Chem. Proc. R&D, Paul.Meenan@dupontpharma.com
Jeff Wolbach, Analytical R&D, jeffrey.p.wolbach@dupontpharma.com; Tel: 302-695-3129, Fax 302-695-9802

BACKGROUND

At the time of this WTEC panel's visit, DuPont Pharmaceuticals was a wholly owned subsidiary of E.I. Du Pont. In June 2001, Bristol-Myers Squibb announced its purchase of Dupont's pharmaceuticals business for $7.8 billion. The transfer was intended to take place by the end of 2001. The company has 5,000 employees, of whom 500 are in research (including about 280 chemists and 220 biologists).

During the last 30 years, DuPont has worked in pharmaceuticals and life sciences; one very successful activity was discovery of the antihypertensive drug Cozaar, the first angiotensin-II receptor antagonist, which was developed and is now marketed for them by Merck. That interaction led to mutual interest and eventually in 1991 to the formation of DuPont-Merck, a 50:50 joint venture in pharmaceuticals. DuPont bought out Merck in 1999, and now fully owns the DuPont Pharmaceuticals Company, a subsidiary of the parent.

DuPont recently acquired Combichem (http://www.combichem.com/), which brings considerable new technology to the company, both in molecular modeling and in experimental combinatorial chemistry. This added about 30 computational chemists. Bridges are being built between the modeling groups, with frequent interactions between Dennis Underwood and Steve Brenner, both at the Experimental Station, and Peter Myers, head of the former Combichem, now DuPont Pharmaceutical Research Labs West Coast, and David Spellmeyer, head of the west coast computational group. There are continual interactions among scientists on the opposite coasts.

R&D ACTIVITIES

Early Activity

DuPont's first computational molecular modeling was in the 1970s when it developed TRIBBLE, an early molecular modeling system (Eaton and Pensak 1978). TRIBBLE eventually evolved into the CAChe product currently marketed by Fujitsu (see separate report). Its calculations included semi-empirical molecular orbital codes such as CNDO and MNDO, the *ab initio* program Hondo and molecular mechanics (MM2). Molecules could be entered by drawing them on a graphics screen. At its peak, 50 to 100 chemists used it.

Other early activity included a collaboration with Columbia University which resulted in FASTRUN, a dedicated computer for molecular modeling developed by Cyrus Levinthal. Although it did not progress beyond a single prototype, it foreshadowed the current IBM "Blue Gene" dedicated protein-folding supercomputer.

In the 1980s, hardware and software developments helped the activity grow. Hardware moved from VAX minicomputers (1981) to the first supercomputer in the chemical industry (a Cray 1-A in 1984) to SGI workstations in the late 1980s. The switch from Tektronix displays to Evans & Sutherland graphical displays in 1983 was another significant step. Meanwhile, use of Amber began in 1981, Chem-X (Oxford Molecular) in 1983, and Biosym and Quanta began in 1988.

Early projects included work on Phospholipase A2, micelle simulations for lipid bilayers, electron transfer in cytochrome proteins, ligand binding for Streptavidin, *ab initio* protein design, and an HIV protease inhibitor sparked by a chemical library search based on a simple pharmacophore.

Present Activities

Carl P. Decicco is an Executive Director in the Chemical and Physical Sciences division, heading a 60-person organization. A synthetic organic chemist by background, he joined DuPont-Merck at its inception. He described two major success stories resulting from combining modeling, X-ray structures and synthetic chemistry in a synergistic way—the development of novel HIV protease inhibitor series and matrix metalloprotease inhibitors. In a 1996 paper (Rockwell et al. 1996), he described in particular the complementary nature of combinatorial chemistry and computational chemistry. Combinatorial chemistry revealed 200 compounds with weak hits. To examine alternatives they tried to dock some of the hits into crystallographic sites, but that did not provide the needed insight. This led to crystallographic analysis of the bound protein, which indicated a conformation change on docking. A patent resulted.

As a consequence, he feels strongly that organic chemists must be able to run visualization codes such as Sybyl and MSI's Insight. He noted that many of his synthetic chemists do spend time in front of the graphics systems thinking about the molecules to make.

The Discovery Informatics and Molecular Design (DIMD) group is headed by Dennis Underwood, formerly of Merck. Organizationally, it is positioned under the direction of Steve Brenner as the head of the Physical Sciences group which resides in the Chemical and Physical Sciences Department. This section includes X-ray crystallography, high-field protein NMR, informatics and modeling, as well as all analytical chemistry. The DIMD group has five people working in discovery cheminformatics, five in molecular design, and three who maintain system software and hardware. This group is expected to work closely with chemists to aid in the discovery of new pharmaceuticals. They also develop new methodologies in data mining and structural genomics. Computing resources include 50 SGI workstations, 10 Sun servers, a 12-processor SGI Origin 2000, and a 16-processor SGI Power Challenge. This hardware is separate from that used for bioinformatics, which resides in a separate Applied Biotechnology Group.

Chong-Hwan Chang described activities in protein crystallography, which are carried out by five PhDs in a group of 15. To supplement their in-house experimental facilities, DuPont co-owns a dedicated beamline at Argonne's Advanced Photon Source, which is rented to DuPont Pharma.

Paul Meenan's work is in process crystallization, where modeling is done mostly with the MSI Polymorph program. He has just moved to DuPont Pharma from DuPont after seven years in crystallization work there. Predictions of shape and structure are both important. While current methods are useful, better ones are needed.

Jeff Wolbach of Analytical R&D described a mix of *ab initio* calculations and information analysis for analytical separations. Calculation of phase equilibrium is consistently an issue for them. Two examples were QSAR analyses for chiral-protein separations by capillary electrophoresis and QSAR for membrane separations.

DISCUSSION

Targets

About two-thirds of the drug discovery projects have something known about the target structure, and half of these have known atomic structures. This has been about constant in recent years, and to some extent most of the industry is examining the same targets. Dr. Underwood felt strongly that these proportions should rise, in particular for targets having partially known structures. High-field NMR, giving fragment information, should give a great increase in the amount of structural information known. The entire structure often is not needed. Among their projects are active programs in modeling and experiment in G-protein-coupled receptors, and they maintain close relationships with others doing research in this area.

Structural information about a target can still be useful late in the development process, but people have to be open to ways they can use it—even if it is for the next project. As one participant pointed out, "Learning a protein's structure is incalculably valuable. Everything changes irreversibly from that point."

Involvement in Intellectual Property

Patents are part of the story, and 50% of the company's discovery method patents include molecular modeling. It is crucial to own the compounds, so any help in developing them is important.

Hardware and Software Issues

Dennis Underwood remarked that in the hardware infrastructure, network stability is critical—the data have to be available in timely fashion. To insure stability, all the SGI workstations are identically configured and all codes are portable. In addition, the modeling group has three Unix people to help insure that stability.

Most of the software is purchased, costing in the very rough neighborhood of $500,000 per year. Pat Brady leads in-house development of new molecular-design codes as they are needed. For informatics, they develop drivers that couple off-the-shelf software. Jeff Wolbach's work uses MSI Cerius^2 mostly. He is not involved in any MSI consortia, although DuPont Pharma is part of the MSI Combinatorial Chemistry consortium.

Culture of Modeling

Modeling here is emphatically to advance development, not "post-dictive" analysis. Each modeler works on two to three projects. He or she is expected to analyze and understand the data and then to exploit it, predict candidates, and rank them. At the same time, flexibility is crucial. Likewise, chemistry and biology still have somewhat different cultures, and the individuals have to go over or around any barriers they find.

Zelda Wasserman noted that in her experience, some of the biggest impacts have used the simplest techniques—short simulations to refine structures, use of simple force fields, and semi-empirical molecular-orbital calculations. This is not to say that bigger, more accurate calculations or projects can't be valuable, but rather that they aren't always necessary to generate significant impact. Counter-intuitive approaches must be recognized, such as examining how a non-prospect binds in order to find better prospects. One speaker re-emphasized that aiding the choice not to go down a path (limiting compound space) is often the biggest contribution.

Steve Brenner noted that there was a spectrum of usage by chemists. However, when a method really works, the bright people are eager to find out how to use it and do so. While it does not seem productive to try and teach everyone how to use molecular modeling tools, it is very important to make learning them possible for people who want to and can justify it.

Handling Interactions Outside the Local Dupont Pharma Site

Interactions include work with DuPont West Coast (Combichem) and with DuPont Corporate R&D. The latter is formally a different company although its common history and proximity (both central R&D locations are at the Wilmington Experimental Station) mean intercompany collaborations are easier. With DuPont Pharmaceutical Research Lab on the West Coast, which is part of DuPont Pharma, interactions are aided by videoconferencing with a dedicated line. Face-to-face interactions are better, but such permanent geographical restrictions require the effective use of various media. The use of Netmeeting technology is being actively explored.

Needed Attributes in Modelers

Sound preparation is necessary, of course; so is circumspection. A person needn't have a wide breadth of specializations, but a breadth of interests and an open attitude are essential. That can be a very different attitude from a student's expectations, so bringing in summer students and post-docs is important.

Promising Future Developments

ADME modeling will play an increasing role as it improves. Identifying a high-affinity ligand is only part of the solution. Its toxicity profile and pharmacokinetic properties must also be optimized, and few computational predictive tools are available in this area.

Although they are not involved in genomics at present, they expect that when interpreted, the genome information will alter styles and directions of research. Knowing that some percentage of the population has a certain trait will even affect the choice of drugs (pharmacogenomics). Also, as Zelda Wasserman commented, genomic data will allow access to new protein structures and provide previously unknown platforms that will be studied even before their functions are known.

REFERENCES

Eaton, D.F., and D.A. Pensak. 1978. *JACS* 100, 7428-7429.

Rockwell et al. 1996. *JACS* 118, 10337-8.

Site: **E. I. du Pont de Nemours and Company**
Central Research and Development
Experimental Station
P.O. Box 80328
Wilmington, DE 19880-0328
http://www.dupont.com/

Date Visited: 20 January 2000

WTEC Attendees: K. Morokuma (report author), P.T. Cummings, R. Miranda, M. Neurock, P. Vashista, R. Wellek, P.R. Westmoreland

Hosts: Dr. Kerwin Dobbs, Central Research and Development, Kerwin.D.Dobbs@usa.dupont.com
Dr. Paul Soper, Central Research and Development, Paul.D.Soper@usa.dupont.com
Dr. John van Stekelenborg, Central Research and Development, John.Van-Stekelenborg@usa.dupont.com
Dr. Rob Waterland, Central Research and Development, Robert.L.Waterland@usa.dupont.com
Dr. Laurie Christianson, Central Research and Development, Laurie.A.Christianson@usa.dupont.com
Dr. Ya-Jun Zheng, Agricultural Products, Ya-Jun.Zheng@usa.dupont.com

BACKGROUND

In 1999, DuPont had revenues of $26.9 billion and 94,000 employees worldwide. Most customers are other manufacturing companies, and its largest markets are in food and nutrition, health care, apparel, home and construction, electronics, and transportation. All the major businesses are based on DuPont inventions.

R&D ACTIVITIES

The R&D budget is ~$1 billion and supports 3700 scientists and engineers. Research is conducted at more than 24 sites; 40% is done in the Wilmington area and 15% (growing) outside the U.S. The Experimental Station, which houses Central Research & Development, is the epicenter of DuPont research; over 40% of R&D is carried out there. DuPont's current technological agenda is to (1) revolutionize asset productivity, (2) double new product introductions, (3) build new polymer platform(s), (4) integrate the biotechnology platform across the company, (5) commercialize new businesses, and (6) develop future opportunities through discovery research.

In order to encourage new technological development, for the first time this year, DuPont initiated a company-wide internal challenge (called Apex) for proposals to discover new areas of science with potential for major business opportunities and to develop new technical solutions for major market opportunities. Each proposal was evaluated for business potential and scientific merit. Of about 500 proposals submitted, about 30 were funded this year, and an additional 90 or so were considered worth funding. Roughly 1-20 (typically 15) persons are involved in each project.

MOLECULAR MODELING ACTIVITIES

The history of the DuPont's molecular modeling effort goes back to the 1970s (see report on the separate company, DuPont Pharmaceuticals). Presently, DuPont's molecular and material modeling efforts include work in a wide range of areas in Central Research and Development and in Agricultural Products:

- Kerwin Dobbs and Paul Soper (Central Research & Development)—Catalysis and Chemical Informatics
  - Dr. Dobbs obtained a PhD in physical chemistry in 1986 with research in computational quantum chemistry, and after two postdoc positions, joined DuPont in 1990. He has since applied computational quantum methods to the investigation of fluorocarbon chemistry, nylon chemistry, and mechanistic pathways of metal-catalyzed homogeneous reactions.
  - Dr. Soper has been at DuPont since receiving his PhD in physical chemistry in 1981. He was originally trained as an experimental spectroscopist, and after a few different assignments, he became a computational chemist, with an emphasis on force fields.
- John van Stekelenborg and Rob Waterland (Central Research & Development)—Chemical process design and manufacturing
  - Dr. van Stekelenborg joined DuPont in 1994 after finishing a PhD in astronomy. He became involved in neural networks, rate constants, kinetics and then computational chemistry, and he carries out mainly quantum chemistry calculations.
  - Dr. Waterland, after a PhD and postdoc in non-linear dynamics, joined DuPont in 1988 and was involved in neural networks and environmental modeling, which led him to computational chemistry of hydrofluorocarbons.
- Laurie Christianson and Elena Kazakova (Central Research & Development)—Molecular diversity analysis & combinatorial library design
  - Dr. Christianson got her PhD in organic chemistry, doing modeling of beta-peptide secondary structure, and joined DuPont in 1997 to do combinatorial library design.
  - Ms. Kazakova joined DuPont in 1998 with a master's degree in materials science. She performs combinatorial library design and QSPR analysis.
- Ya-Jun Zheng, Dan Kleier, John Kinney, Kevin Kranis, Bill Tillotson, and Mei Dung (Agricultural Products)—Crop protection chemical discovery
  - Dr. Zheng got a PhD in organic chemistry and, after a postdoc, came to DuPont in 1997 to carry out quantum mechanical calculations, protein simulations, and structure-based molecular design.
  - Dr. Kleier joined DuPont in 1986 as a molecular modeler and is the technical group leader of the computational chemistry group in Crop Protection Products. He has a PhD in physical organic chemistry.
  - Dr. Kinney has been at DuPont since 1981 after completing a PhD in physical chemistry. He applies methods of statistical and chemical diversity analysis to compound acquisition and library design.
  - Mr. Kranis, who specializes in data analysis and library design, joined DuPont in 1987 with a bachelors degree in chemistry.
  - Dr. Tillotson has a PhD degree in soil physics. Since joining DuPont in 1992, he has modeled the behavior of chemicals in soil/plant systems and performed structure-activity analysis in support of lead optimization.
  - Dr. Dung began working at DuPont in 1998 with a PhD in physical chemistry. She is presently doing molecular modeling in support of lead optimization.

The main computer hardware is composed of two SGI 22-CPU Origin 2000 machines located at the Experimental Station site. The computationally demanding software programs are run in two- or four-fold parallel mode, using a batch queuing environment which evenly balances the load. Major software packages used for quantum chemistry modeling are Gaussian with GaussView and DGauss with UniChem; in addition, Jaguar (for pK_A in solution; Schrödinger, Inc.), Tripos Sybyl, and MSI codes are used. Isodesmic reaction techniques as well as CBS, MP2, and B3LYP levels of calculations are used for energetics and kinetics. Rapid analysis of a large chemical library is performed on their high performance computers using codes such as Tripos' Concord and Diverse Solutions as well as MSI's combinatorial chemistry software. Modeling is also done for polymers; for instance, photodecomposition of polymers using small oligomers. No mesoscale modeling is performed. Code development is currently restricted to that of database capability.

Work of the computational chemists is sometimes recognized in patents. For instance, some patents in catalysis for hydrofluorocarbon degradation and in crop protection chemicals involve computational chemists. The head of the agricultural computational chemistry group, Dan Kleier, won a "scientific leadership" award for his contribution.

The primary goal of computational chemistry is to improve research productivity. The computation should be reliable (to extend limited experimental data and study systems confidently which are difficult experimentally), cost-effective (to guide and screen experiments), and mutually stimulating between experimentalists and computational chemists (to lead to new models and concepts and to new insight). The second goal is to provide a competitive edge (used as a strategic tool). They try to integrate computational chemistry technology into key focusing problems. It is recognized that the best use of computational chemistry is in the beginning stages of long-term programs (such as Apex); it develops synergism between computational and experimental researchers. Also, early completion of validation work saves time later in the program. Computational chemistry is a reusable resource, both for "soft" knowledge (understanding, concepts models) and for "hard" knowledge (storable in a database). Apex is expected to help computational chemistry.

Validation of the computational results is considered to be paramount. "How well do the calculations agree with what is known by observation and experiment?" There is a need to understand predictive capabilities, both qualitative and quantitative.

Molecular thermochemistry is one of the important applications of computational chemistry at DuPont. Thermodynamic properties (ΔH, ΔS, ΔG, C_v, C_p, ΔH_f, reaction energies, bond strengths, binding energies and activation energies) need to be known within 1-5 kcal/mol. Computed gas-phase results can be approximately compared with the experimental results in non-polar solvents. Solvent effects are currently ignored but may need to be investigated in the near future. Computational thermochemistry has demonstrated value in safety awareness, early scouting of novel chemistries, and providing key data for process models. It is also used to optimize design of chemical processes.

Computational support of combinatorial chemistry research requires integration of traditional computational chemistry with chemical information. The scientists working in this area align their efforts with those of programmers and database designers to allow better access to the large amounts of data combinatorial chemistry generates. The Central Research and Development combinatorial chemistry group focuses on non-biological applications. Traditional molecular diversity methods are often only partially applicable to non-bioactive compounds, and methods development and validation are active areas of interest and research.

Other views on the status of computational chemistry are as follows:

- "Consortiums are dead!" Some are too focused and too many companies are involved and often miss the points. Those that do provide value usually require constant vigilance on the part of members.
- Adopt a "fast-follower strategy," i.e., maintain computational chemistry competence so DuPont can benefit from advances without large development cost.
- Development of computational chemistry is still in its "childhood." It is not yet the "commodity item" anybody can use. Nevertheless, DuPont recognizes use by its competitors, including Air Products, BP-Amoco, BASF, Dow/Union Carbide, DSM, Exxon, Goodrich, Hoechst, Phillips, and Shell.
- For value recognition, computational chemistry needs long-lasting, highly successful "hits." The hydrofluorocarbon project to replace chlorofluorocarbons (CFCs) was a major success story at DuPont.

REFERENCES

CR&D References

Coburn, J.C., P.D. Soper, and B.C. Auman. 1995. Relaxation behavior of polyimides based on 2,2'-disubstituted benzidines. *Macromolecules* 28, 3253.

Crawford, M.K., K.D. Dobbs, R.J. Smalley, D.R. Corbin, N. Maliszewskyj, T.J. Udovic, R.R. Cavanagh, J.J. Rush, and C.P. Grey. 1999. A raman spectroscopy study of the separation of hydrofluorocarbons using zeolites. *J. Phys. Chem.* 103, 141.

Dixon, D.A., A.J. Arduengo, III, and K.D. Dobbs. 1995. On the proposed existence of a ketene derived from carbon monoxide and 1,3-di-1-adamantylimidazol-2-ylidene. *Tetrahedron Lett.*, 36, 645.

Dixon, D.A., and K.D. Dobbs. 1994. Computational chemistry for the design of materials with low environmental impact. *Trans. Mater. Res. Soc. Jpn.*, 18A(Ecomaterials), 373.

Dixon, D.A., K.D. Dobbs, and S.C. Walker. 1994. Computational chemical design for CFC alternatives at DuPont. *Cray Channels*, 16, 7.

Dixon, D.A., K.D. Dobbs, M. Neurock, J. Lerou, and T. Nakao. 1995. The prediction of molecular properties for product life cycles. *Proc. 9th Int. Symp. Large Chem. Plants*, Antwerp, Belgium, p 101.

Dobbs, K.D., and D.A. Dixon. 1994. Ab initio prediction of the activation energy for the abstraction of a hydrogen atom from methane by a chlorine atom. *J. Phys. Chem.*, 98, 12584.

Dobbs, K.D., and D.A. Dixon. 1996. Reliable reaction enthalpies for neutral amide hydrolysis and the heat of formation for formamide. *J. Phys. Chem.* 100, 3965.

Dobbs, K.D., D.A. Dixon, and J.J. Valentini. 1994. Amide-water and amide-amide hydrogen bond strengths. *J. Phys. Chem.*, 98, 13435.

Saitta, A.M., P.D. Soper, E. Wasserman, and M.L. Klein. 1999. Influence of a knot on the strength of a polymer strand. *Nature* 399, 46.

Agricultural Products References

Chen, J.M., S.L. Xu, Z. Wawrzak, G.S. Basarab, D.B. Jordan. 1998. Structure-based design of potent inhibitors of scytalone dehydratase: Displacement of a water molecule. *Biochemistry* 37, 17735.

Jordan, D.B., Y.-J. Zheng, B.A. Lockett, and G.S. Basarab. 2000. Stereochemistry of the enolization of scytalone by scytalone dehydratse. *Biochemistry* 39, 2276.

Kleier, D.A. 1993. Quantitative structure-activity relationships: Adding value to herbicide bioassays. In *Herbicide Bioassays*, J.C. Streibig & P. Kudsk (Eds.), CRC Press, Boca Raton, FL., pp. 97-110.

Kleier, D.A. 1994. Environmental effects on the photodegradation of pesticides. In *Comparing Glasshouse & Field Pesticide Performance* (H. G. Hewitt, J. Casely, L. Copping, and B. T. Grayson, Eds.), BCPC Monograph 59; BCPC Publications, Farnham, Surrey, UK, pp.97-110.

Kleier, D.A. Molecular modeling and simulation of crop protection chemicals. In *Foundations of Molecular Modeling and Simulations* (P.T. Cummings, P.R. Westmoreland, and U.W. Suter, Eds.), AIChE Symposium Series No. 325, Volume 93 (in press).

Zheng, Y.-J., and T.C. Bruice. 1998. Role of a critical water in scytalone dehydratase-catalyzed reaction. *Proc. Natl. Acad. Sci. U. S. A.* 95, 4158.

Zheng, Y.-J., P.V. Viitanen, and D.B. Jordan. 2000. Rate limitations in the lumazine synthase mechanism. *Bioorganic Chemistry* 27, 89.

Zheng, Y.-J., R. Shapiro, W.J. Marshal, and D.B. Jordan. 2000. Synthesis and structural analysis of the active enantiomer of famoxadone, a potent inhibitor of cytochrome bc1. *Bioorganic & Medicinal Chemistry Letters* 10, 1059.

Site: **Eastman Chemical Company**
P.O. Box 1972
Kingsport, TN 37662-5150
http://www.eastman.com/

Interview Dates: December 27, 1999, and January 3, 2000

Interviewer: P.R. Westmoreland

Contacts: Peter M. Margl (presently at Dow Chemical), J. I. Siirola

BACKGROUND

Eastman Chemical (http://www.eastman.com/), the former chemicals arm of Eastman Kodak, has a large manufacturing complex in northeastern Tennessee. Among its other accomplishments, it is the only company in the U.S. that uses coal as a major chemicals feedstock.

There were two fulltime persons, Peter Margl and Bruce Wilson. Margl's background involved DFT of homogeneous catalysis with Tom Ziegler (Calgary), while Wilson's PhD in analytical chemistry (Univ. Washington) focused on chemometrics. There also were a few bench chemists who did modeling on personal computers with Spartan (Wavefunction, Inc.) and Titan (Schrödinger, Inc.).

The computational chemistry group was part of a larger laboratory that was otherwise not involved with computer simulation. That laboratory was part of the Chemicals Research Division. The reason for placing the modelers in this structure was that for a group of two people, the organizational overhead in forming a new laboratory would have been prohibitive. In late 1999, the decision was made to refocus research and development, which included termination of the computational chemistry effort.

MOLECULAR MODELING R&D ACTIVITIES

Work in molecular modeling was started in 1997 with the attempt to model epoxybutene catalysis (heterogeneous process, catalyzed by a promoted silver surface). Attempts were made to model the process with periodic boundary conditions and in a cluster model.

Among the topics of interest were the following:

- Heterogeneous catalysis (both on metals and insulators)
- Homogeneous organometallic catalysis (polyolefin and hydroformylation)
- General thermochemistry (organic and inorganic)
- Organic reaction mechanisms (thermodynamics and kinetics)
- Process kinetics and thermodynamics

DISCUSSION

The two modeling specialists usually worked in such a way that one person was working on one project, with minimal overlap. Each modeler was responsible for maintaining contact with internal customers and for selecting worthwhile projects. Margl and Wilson got together practically every day to discuss strategies and further developments.

Information management is necessarily important. The modelers developed a platform for information exchange and storage within the computational chemistry group, as well as for communication of results to contacts outside the group. This tool was based on Microsoft Access and was basically a database with a visual interface that could take over job control and job management to some degree.

Method validation was usually done by comparison with high-quality experimental data for model chemistries that were as close as possible to the process chemistry. If experimental data were absent, the highest possible level of theory was employed for the model chemistry. Also, consistency of the results with respect to basis set size and level of theory was checked.

The Eastman modelers found that computational chemistry had accumulated a rather bad reputation in the last twenty years, which it had trouble shedding. They perceived this also as a wider issue throughout industry. This seemed to be because, when computational chemistry first became popular, it was not much of a practical day-to-day tool, and its specialists felt that they had to massively oversell in order to maintain their jobs. Now that computational chemistry is becoming really practical, many chemists have still preferred to stay away from it ("once bitten, twice shy"). However, many people, especially those who do modeling on their own PCs, began to see the applicability of modeling right away. Generally, Margl and Wilson felt they usually had to undersell themselves in order to avoid any suspicion of overselling. Consequently, their strategy for dealing with customers was to outline the capabilities of modeling very clearly, along with what could be expected of it in the context of the customer's problem right at the start. They also strove to give as realistic error bars as possible.

Of course, failures did occur at times. They considered their failures as coming mainly from two sources: modeler errors and inadequacies of the theory they used. They strove to get rid of the former by letting other people think through the modeling strategy. They tried to eliminate errors coming from the latter by rigorously checking against experimental data. Of course, this may mean that it turns out that the problem is not solvable by current computational techniques. That particularly pertains to problems involving solvation, systems with too many degrees of freedom and systems with tricky transition metal centers, especially if there is more than one transition metal involved. Interestingly, they usually found that computational chemistry was up to most challenges.

A particular wish was for a better handle on explicit solvation and transition metal systems. Modern DFT is doing a good job on most, but when there is a system with many almost-degenerate states, DFT is not applicable. A practicable QM/MM methodology that allows for some amount of charge transfer between QM and MM sides would be valuable.

Site: **Ford Motor Company**
The American Road
Dearborn, MI 48121
http://www.ford.com/

Date of report: April 2001

Author and contact: Ellen B. Stechel, Manager of the Chemistry & Environmental Science Department, Ford Research Laboratory [and member of this panel]

BACKGROUND

Ford Motor Company is the largest manufacturer of trucks and the second largest manufacturer of cars and trucks combined in the world. It includes under its corporate umbrella many of the most familiar vehicle brands in the world, including Ford, Lincoln, Mercury, Volvo, Mazda, Jaguar, Land Rover, and Aston Martin. The company is also one of the largest providers of financial services worldwide, including financing, vehicle and equipment leasing and rental operations (including Hertz), and insurance operations. It is number two on the Fortune 500 list of largest companies in the world, with annual revenues in 1999 of over $160 billion and net income of over $7 billion.

RESEARCH ACTIVITIES

Overview

Ford Motor Company research activities are located in the Ford Research Laboratory (FRL), which includes laboratory facilities in Dearborn, Michigan and Aachen, Germany. FRL is a part of the global product development organization and performs advanced research related to all aspects of the Ford automotive business. Its mission is to "anticipate the technical needs of the company and its customers," "innovate solutions to technical challenges," and "incorporate developed technology into products and processes."

The Chemical and Physical Sciences Laboratory (C&PSL) within FRL includes research in materials science, environmental catalysis and emissions control, environmental science, and physics, and is staffed with approximately 200 scientists and engineers. Despite the fact that today Ford produces few of the materials used to build new vehicles, the company maintains active research programs in many areas of chemistry and materials science, including polymers and composites, paints, alloys, and catalysts. These activities serve essentially three purposes: first, to develop new materials, processes, and technologies that will meet evolving consumer and market demands and distinguish the company's products in the marketplace; second, to make current materials, processes, and technologies more dependable and robust; and third, to interface with suppliers to ensure that their current and future products will meet the company's needs. Environmental drivers, including decreasing regulated and non-regulated emissions and improving the fuel economy of vehicles, motivate much of this research. As a result, a significant research effort in C&PSL is also directed at understanding and quantifying the environmental impact of vehicles, especially on regional atmospheres and the climate.

History of Atomistic and Molecular Modeling Research

Molecular modeling research was conducted within the predecessor organizations of FRL as long ago as the early 1970s. The current effort began in the early 1990s, when opportunities appeared to apply molecular modeling to atmospheric chemistry—specifically the atmospheric degradation of alternative refrigerants chosen to replace chlorofluorocarbons—and to catalysis for emissions control—specifically NO_x reduction catalysis required to meet increasingly stringent emissions control standards. Each of these began as "grass roots" efforts, initiated by different researchers who saw opportunities to apply molecular modeling to existing research projects. The scientific successes of both of these projects led to the development of a more coordinated, although still organizationally separated, program, and to extensions into other areas of application, including adhesion, photochemistry, and combustion. The success also led to the gradual growth

in program size, from an initial staffing of two researchers to the current level of six. As the program has grown, so has the range of problems explored, so that now "molecular modeling" within Ford is seen broadly to encompass all types of first principles, atomistic, and mesoscale simulations.

The molecular modeling activity is divided across chemistry and physics within the C&PSL organization. This decentralization exists in part for historic reasons, but is seen to have the advantage of promoting a greater set of interactions with experimental research programs. The "modelers" typically are not assigned to a specific project team either, but rather are given considerable latitude to choose one or more projects to support based on their assessment of the probability for impact.

The long-range vision of this "virtual group" is to grow molecular modeling into an integral part of the research and development processes at Ford, and in doing so to reduce the development time and cost, and to improve the features and quality, of the company's products. Its present mission is to complement and eventually to guide experimental research programs, to accelerate the development and selection of new technologies for Ford products. Because of the great complexity of the chemical and materials challenges faced by Ford, molecular modeling has had its greatest successes in generating fundamental scientific insight and understanding as opposed to direct solution to technological problems.

The group is held to high scientific standards, publishes its research in first-tier scientific journals, and is expected to regularly participate in and contribute to scientific conferences. It also uses external academic collaborations and graduate student internships extensively to leverage its work.

Recent Activities

- *NO_x SCR catalysis.* Lean-burn gasoline and diesel engines offer the promise of 10 – 15% improvements in fuel economy. Because the exhaust from these engines is rich in O_2 and poor in reductants, traditional "three-way" catalysts (alumina-supported noble metals) are ineffective in removing NO_x from the exhaust. Cu-exchanged zeolite ZSM-5 has amongst the highest known activities for this selective catalytic reduction (SCR) of NO_x. Density functional theory simulations have been used to probe NO_x chemistry in these materials, with the expectation that improved understanding will lead to the development of more active and more robust catalysts. Cluster calculations have been used to determine the structure of active sites (Schneider et al. 1996; Goodman et al. 1998), to probe the vibrational spectroscopy of adsorbates (Ramprasad et al. 1997), and to elucidate mechanisms of NO_x decomposition and reduction chemistry (Schneider et al. 1997; Schneider et al. 1998a). Much of this work was carried out with collaborators at Arizona State University, and in particular, with graduate students who spent their summers at Ford working on the project.
- *Molecular tribology.* Molecular tribology involves using molecule-scale theories, simulations, and experiments to provide in-situ, molecular-level descriptions of lubricant and lubricant additive actions between sliding surfaces. Classical molecular dynamics (MD) simulations have been used at Ford to study the so-called "friction modifier" additive found in automatic transmission fluid. This additive adsorbs on clutch surfaces and reduces both static and low-speed dynamic friction, leading to a qualitative change in the low-speed dynamic friction coefficient: increases in friction with velocity are observed, compared to the decreases found in base oils. Simulation results to date (Greenfield and Ohtani 1999) have helped explain experimental data (Ruths et al. 1999) measured with a surface forces apparatus (SFA) by Ford researchers (in conjunction with the University of Illinois at Urbana-Champaign) on how the separation between surfaces changes with imposed pressure, in the absence of sliding. In ongoing work, the velocity dependence of the friction coefficient is being studied through both simulations (MD) and experiment (SFA). The long-term objective is to determine how molecular factors (lubricant size, shape, energetics, etc.) affect sliding friction in both boundary and elasto-hydrodynamic lubrication of surface–fluid–surface contacts.
- *Alumina structures and surface chemistry.* Aluminas are of considerable technological significance, for instance as supports and active components in catalysts and as interfaces in adhesively bonded aluminum. Because of their importance, a variety of projects have been or are addressing alumina structure and chemistry. In collaboration with IBM-Zürich, Car-Parrinello molecular dynamics (CPMD) simulations were used to study the chemistry of water on an alpha-alumina surface (Hass et al. 1998). In addition to addressing an important scientific problem, this collaboration was used as a mechanism to

gain experience with the CPMD approach and with parallel computing on the IBM SP platform. These first principles calculations complemented collaborative work with Wayne State University to use empirical potential molecular dynamics to model water and alkanes on the same surface (de Sainte Claire et al. 1997; Bolton et al. 1999), and together form a basis for beginning to understand the competition between adhesives and adventitious adsorbates in bonding to oxide surfaces. The bulk and surface structures of gamma-alumina are considerably more complicated than that of the alpha phase: DFT calculations are being used to model the bulk structure of gamma-alumina, and in particular to unravel the role of hydrogen in stabilizing that phase.

- *Al alloy castings.* Increasing pressures to further reduce emissions and simultaneously increase fuel economy in automobiles has created a need for lightweight materials (such as Al, Mg, and their alloys). In order to optimize alloy design and processing conditions to achieve Al-alloy castings with the necessary mechanical properties, researchers at Ford Research Laboratory are developing the "virtual aluminum castings" methodology: a suite of predictive computational tools that span length scales from atomistic to macroscopic to describe alloy microstructure, precipitation, solidification, and ultimately, mechanical properties. First-principles atomistic computations have recently been added to the set of methodologies in this project. Previously, micro-structural issues in alloys, such as precipitation, have largely been outside the realm of first-principles atomistic calculations. However, the development of a new method enables the prediction of coherent phase equilibria and precipitate shapes in Al-alloys with system sizes up to 250,000 atoms (Wolverton 1999). These types of calculations may then be combined with larger-length-scale micro-structural models, and in conjunction with experimental efforts, can be used to suggest heat treatments that optimize thermal stability and hardness of industrial alloys.
- *Oxides for sensors, fuel cells, and oxygen storage.* Rare earth oxides are used as oxygen storage components in automotive catalyst formulations and as oxygen transport components in high temperature sensors and fuel cells. An effort has recently been initiated to use DFT methods and kinetic Monte Carlo to model ionic transport in doped zirconias.
- *NO_x trapping catalysts.* An alternative to NO_x SCR is the selective trapping of NO_x from automotive exhaust during lean burn operation and purging and reduction of NO_x during periodic rich excursions. These NO_x trapping catalysts are typically composed of an alumina support with dispersed alkaline earth oxides to serve as trapping materials and dispersed noble metals to catalyze oxidation (during lean trapping) and reduction (during purging) reactions. The current formulations are rapidly poisoned by fuel sulfur, which is the primary obstacle to adopting this technology. DFT calculations are being used to study the NO_x and SO_x trapping reactions on the alkaline earths, with the goal of designing formulations with high activity for NO_x adsorption and simultaneously high sulfur tolerance. A collaboration with MIT focuses on the role of the noble metals in catalyzing oxidation and reduction reactions.
- *Atmospheric and combustion chemistry.* FRL has a strong effort in experimental atmospheric chemistry, and molecular modeling has frequently been used to support this activity, for instance through explanation of observed reactivity patterns (Schneider et al. 1998b) and through calculation of thermodynamic properties (Schneider and Wallington 1994). A current collaboration with UCLA is using *ab initio* calculations and molecular dynamics to explore the low-temperature combustion chemistry of dimethyl ether, a promising alternative diesel fuel.

Personnel

- *Dr. Kenneth C. Hass* has been with FRL in the Physics Department for twelve years. His background is in condensed matter physics. He initiated the molecular modeling program in NO_x catalysis and continues to work in that area, as well as on metal oxide structure and surface chemistry.
- *Dr. William F. Schneider* has been with FRL in the Chemistry Department for eight years. His background is in inorganic chemistry. He initiated the molecular modeling program in atmospheric chemistry and continues to contribute in that area. His current emphases are NO_x catalysis and oxides for NO_x traps.
- *Dr. Michael L. Greenfield* has been with FRL in the Chemistry Department for four years. His background is in chemical engineering and polymer statistical mechanics, and he leads the research effort in tribology.

- *Dr. Christopher M. Wolverton* has been with FRL in the Physics Department for one year. His background is in Solid State Physics and Materials Science, and he leads the research effort in Al alloy microstructure and is involved in modeling oxides.
- *Dr. Alexander Bogicevic* has been with FRL in the Chemistry Department for four months. His background is in Physics and Civil Engineering, and he leads the research effort in high-temperature oxygen transport materials.
- *Mr. David J. Mann* has just recently joined FRL in the Chemistry Department. His background is in physical chemistry, and he is responsible for maintaining group computer and software resources as well as supporting various research programs.

Software Resources

The group uses a variety of academic, commercial, and in-house codes to accomplish its research. Density functional theory (DFT) software is a workhorse for many projects. Both local orbital codes—including ADF, DMol, and Mulliken—and plane-wave codes—such as VASP, CPMD, and the all-electron FLAPW method—are or have been used for various projects. Both the GAMESS and ACESII codes are available for wave-function-based *ab initio* calculations, and the LAMMPS code from Sandia is used for force-field-based molecular dynamics simulations. More specialized codes are also used for kinetic and statistical mechanical calculations. Ford purchased Cerius2 and a variety of supporting modules from MSI several years ago, primarily for its visualization capabilities, and maintains licenses on this software. Ford does not participate in the MSI consortia.

Computational Resources

Like the group itself, its computational resources are not centrally managed. Every researcher has an up-to-date Unix workstation available for personal use in addition to a standard issue personal computer. In addition, FRL maintains a 64 processor Origin 2000, which is shared with other projects, while the group maintains a dedicated 16 processor IBM SP. The group is exploring migrating some of its efforts to Linux-based platforms.

CONCLUSIONS

The Ford molecular modeling team is quite strong and active in a variety of areas of technological significance. It has undergone recent rapid growth, reflecting confidence in the future prospects for positively impacting technology development at Ford.

REFERENCES

Bolton, K., S.B.M. Bosio, W.L. Hase, W.F. Schneider, and K.C. Hass. 1999. *J. Phys. Chem. B* 103, 3885-3895.

de Sainte Claire, P., K.C. Hass, W.F. Schneider, and W.L. Hase. *J. Chem. Phys.* 106, 7331-7342.

Goodman, B.R., W.F. Schneider, K.C. Hass, and J.B. Adams. 1998. *Catal. Lett.* 56, 183-188.

Greenfield, M.L., and H. Ohtani. 1999. *Tribo. Lett.* 7, 137-145.

Hass, K.C., W.F. Schneider, A. Curioni, and W. Andreoni. 1998. *Science* 282, 265-268.

Ramprasad, R., W.F. Schneider, K.C. Hass, and J.B. Adams. 1997. *J.Phys. Chem. B* 101, 1940-1949.

Ruths, M., H. Ohtani, M.L. Greenfield, and S. Granick. 1999. *Tribo. Lett.* 6, 207-214.

Schneider, W.F., and T.J. Wallington. 1994. *J. Phys. Chem.* 98, 7448-7451.

Schneider, W.F., K.C. Hass, R. Ramprasad, and J.B. Adams. 1996. *J.Phys.Chem.* 100, 6032-6046.

Schneider, W.F., K.C. Hass, R. Ramprasad, and J.B. Adams. 1997. *J. Phys. Chem. B* 101, 4353-4357.

Schneider, W.F., K.C. Hass, R. Ramprasad, and J.B. Adams. 1998a. *J. Phys. Chem. B* 102, 3692.

Schneider, W.F., T.J. Wallington, J.R. Barker, and E.A. Stahlberg. 1998b. *Ber. Bunsensges. Phys. Chem.* 102, 1850-1856.

Wolverton, C. 1999. *Phil. Mag. Lett.* 79, 683.

Site: **General Electric Corporation**
http://www.ge.com/

Interview Dates: Feb. 7, 2000 (Whitney); Sept. 30, 2000 (Fontana)

Interviewer: P.R. Westmoreland

Contacts: John M. Whitney, John.Whitney@crd.ge.com
GE Corporate Research & Development
K1-5A10, One Research Circle, Niskayuna, NY 12309
518 387-6847 phone, 518 387-5592 FAX
Luca Fontana, luca.fontana@gepex.ge.com
Manager, Business Development and e-Business
GE Plastics, Global Technology
Pittsfield, MA
413-448-4662 phone, 413-448-7635 FAX

BACKGROUND

General Electric (http://www.ge.com/) is an international company with a wide range of business sectors, including polymers, appliances, lighting, electronics, adhesives and sealants, and financial services. It employs over 340,000 employees in more than 100 countries, including 197,000 employees in the United States. Its 1999 revenues were $112 billion, and its market capitalization of $500 billion is second largest in the world.

MOLECULAR MODELING R&D ACTIVITIES

Work in molecular modeling at GE has only recently resumed. In the late 1980s and early 1990s, Dr. John T. Bendler led a well-funded program in modeling polycarbonate properties, focusing mainly on bisphenol-A polycarbonate oligomers. That work used commercial codes from software companies that subsequently grew into MSI. In the early 1990s, GE Plastics decided to refocus its efforts toward the costs of products and processes, soon followed by the Six-Sigma quality initiative. Bendler's activity was terminated at that time.

In the last three to four years, there has been renewed interest in molecular modeling for polymer properties. The polymer development groups in Schenectady, NY and Mt. Vernon, IN have begun to refocus on new copolymers, especially in the polycarbonate and polyester areas. They have specific interest in resin that could meet the requirements of next-generation media applications as well as weatherable resins for molded-in color applications (mainly automotive).

Dr. Whitney is the one person presently working at GE Corporate R&D in this area, although a counterpart will soon be added at the company's India Technology Center. The general area of application is properties of polymers, mainly polycarbonates to be used for CDs, DVDs, and barrier coatings.

At present or in the near term, codes being used on PCs are Spartan (Wavefunction, Inc.), SciPolymer, Alchemy2000 (Tripos), SciQSAR, Polymer Design Tools, and some ACD software. On the company's supercomputer, Spartan is used, available to others within the company by X-Windows. Also important are in-house codes to model behavior of parts, using basic polymer properties and continuum models.

DISCUSSION

The renewed interest in molecular modeling is spearheaded by a market-driven need that shows the limitation of existing technologies. It is strengthened by two critical factors:

1. The addition of R&D resources through the creation of the John F. Welch Technology Center in Bangalore, India.

2. The availability of Web-based collaborative tools that will facilitate the exchange of data and models among the various GE sites around the world. GE has adopted new software for data management/sharing, eMatrix, to enable the digitization of its R&D and engineering processes.

Molecular modeling, coupled with combinatorial chemistry efforts, must demonstrate the speed advantage that is needed to accelerate the new-product introduction cycle. Management's assessment is simply that molecular modeling had yet to yield significant successes for GE.

Dr. Whitney's opinion is that interest by individual researchers is shaded by cautiousness because of past negative experiences. Partly as a result, resistant attitudes arise among longtime synthetic organic chemists. Some of the more experienced people feel that their existing chemical intuition is adequate, so that chemical modeling is neither needed nor wanted. Another such attitude is simple reluctance to change. Again, all parties recognize that internal successes would stimulate stronger interest.

Nevertheless, there is commitment to make use of these approaches when they can be helpful. Dr. Whitney envisions a "Virtual GE Laboratory" in which chemists at GE can go from idea generation up to virtual product/part fabrication in a few days. Not only modeling but also information technology is required, not all of which is available or integrated at present. For example, such a virtual laboratory requires seamless integration of database searching/management (CAS or in-house), small molecule molecular modeling (monomers, catalysts, additives, virtual characterization of these, and true-to-life reaction simulation), macromolecule modeling (material properties, polymer properties, and true-to-life polymerization simulation), and part/product modeling (molding, part properties, cost, etc).

Reaching these goals will be aided by the rapidly changing hardware, software, and attitudes. PCs are the perceived choice for hardware as they continue to be faster and cheaper. Density-functional routines will become quicker, and many information technologies are developing that will be useful. Still, more vendors must offer complete suites of tools for chemists (e.g., ACD and MSI).

Site: **HRL Laboratories, LLC**
3011 Malibu Canyon Road
Malibu, CA 90265
http://www.hrl.com/

Interview date: 5 September 2000

Interviewer: P.R. Westmoreland

Interviewee: Dr. Mark Gyure, gyure@hrl.com

BACKGROUND

HRL Laboratories, formerly Hughes Research Laboratories, was founded in 1948 as the Electronics and Guided Missile Department of Hughes Aircraft Company. It was renamed in 1997 and changed to 50/50 ownership by Hughes Electronics and Raytheon. Presently [2001], it is owned jointly by Raytheon, Boeing, and General Motors, providing corporate research and development for all three companies.

HRL's focus is applied research in electronics, photonics, and information sciences. Target activities include developing products, processes, and services for space, telecommunications, defense and automotive applications. To these ends, it is organized into four technical laboratories: Communications & Photonics, Information Sciences, Sensors & Materials, and Microelectronics. Molecularly based materials modeling at HRL is located in the Computational Physics Department, part of the Information Sciences Laboratory (http://www.hrl.com/TECHLABS/isl/isl_home.shtml).

MOLECULAR MODELING R&D ACTIVITIES

Mark Gyure (PhD Physics, Colorado; post-doc Boston University) and Richard Ross (PhD Univ. of California at Santa Barbara) lead the molecularly-based materials modeling at HRL. They are currently joined by a post-doc and a PhD candidate, Geoff Simms, who is an HRL employee completing his PhD at the University of Illinois.

Such modeling began at HRL in 1994 after Dr. Gyure was hired. HRL was looking to expand its expertise in computational physics, and materials physics seemed to be a promising area. HRL researchers became more specifically interested in modeling III-V compound semiconductors. Both for military and high-end commercial applications, they became interested in processes and properties for nanolayered materials made by molecular-beam epitaxy. The reaction chemistry of this deposition is relatively straightforward, but the interfaces between materials must be precisely maintained for a high level of quality.

To these ends, they have applied both molecular and continuum modeling. They developed their own kinetic Monte Carlo code for the growth process. They also use the *ab initio* molecular dynamics code FHImd-98 of Scheffler and co-workers at the Fritz-Haber Institute of Berlin. Their collaboration with Scheffler began with hiring a knowledgeable post-doc from Scheffler's group, and it has been maintained by visits to the Fritz-Haber Institute. Continuum modeling is carried out with an applied math group from UCLA with DARPA funding.

Computer resources include UNIX workstations, where most of the work is carried out; a 44-node Linux cluster of 650 MHz Pentium IIIs; and an eight-processor SGI Origin 2000 with 8 gigabytes of memory. They have not used MSI software, in part because of the cost of long-term leasing arrangements.

DISCUSSION

To be most effective, modelers must build and maintain credibility internally and externally. HRL recognizes the importance of having a core of people doing top-flight, frontier work to be able to communicate with other researchers at the frontier. Internally, it is important to deliver results—"even if not

indispensable, [the results] must at least be useful." Likewise, it is important that managers be kept accurately briefed on the capabilities of the modeling, both in terms of power and of limitations.

Among internal clients, at present there seem to be three levels of interest in the theory group's tools. First are the people with no interest; happily, few such people are at HRL. Second are people who are interested only if modeling can answer a specific question, finding it useless if not. Third is the group of people who are interested and open to the possibilities. Give-and-take discussions with such people have proven very productive.

Close coupling to experiment is important. Being able to "talk in both languages" is a requirement.

As new methods are used, method validation becomes invaluable. This includes comparisons to previous calculations as well as to experiments. At the same time, it is important to make comparisons recognizing the details of experimental uncertainties. When this is done, trends and idealized predictions can drive development.

Collaborations have been effective, but collaboration at a distance has also had its disappointments. The first big effort was a strongly motivated project with collaborators from Harvard, Colorado, UC Santa Barbara, and UCLA. Portions of this collaboration worked extremely well, but it proved hard to work together closely enough with everyone and have enough interactions. Occasional face-to-face meetings were helpful, including having post-docs come to HRL for extended stays (months). HRL staff members concluded that a good rule of thumb was "ten e-mails equal five minutes at a blackboard."

For HRL's interest in computational materials science, reaching large-sized calculations is important. That places high demands on computers and codes. *Ab initio* MD development is a key aspect, and they have worked effectively with the Fritz-Haber Institute, which is a leader in this area. Order-N methods are also of interest. However, HRL researchers could have a greater impact simply by being able to do today's size of problems faster.

Site: **Louisiana State University**
Concurrent Computing Laboratory for Materials Simulations
Baton Rouge, LA 70808
225-388-1157; Fax: 225-388-5855
http://www.cclms.lsu.edu/

WTEC Author: P.R. Westmoreland

Technical source: Professor Priya Vashishta

BACKGROUND

The Concurrent Computing Laboratory for Materials Simulations was formed in 1992 when Louisiana State University brought Priya Vashishta (a member of this WTEC panel), Rajev Kalia, Simon de Leuew, and Akiro Nakano from Argonne National Laboratory into the departments of Physics and Computer Science. De Leuew later moved to Delft (http://www.cp.tn.tudelft.nl/home/deleeuw/), but the rest lead a team of scientists focused on very-large-scale computing for condensed-matter physics.

The CCLMS claims two main objectives: (1) multidisciplinary research involving grand challenges in materials simulations, and (2) the development of undergraduate and graduate education in computational sciences. Toward the latter objective, it has created a dual degree program consisting of a PhD in physical sciences and MS in computer science.

Hardware in the center includes two parallel computing laboratories including a large Beowulf cluster. The largest calculations are carried out on DoD supercomputers as part of the DoD Grand Challenge Awards.

MODELING R&D ACTIVITIES

Examples of CCLMS modeling activities are presented below. The topics represent simulations of optoelectronics structure (a-GaAs), oxidation and sintering of ceramic nanoclusters, crack propagation in ceramics, nanoindentation, and ceramic fiber composites.

Structure of Amorphous Gallium Arsenide for Optoelectronics

Amorphous GaAs is an important material in various optoelectronic applications. Commonly produced by ion implantation of crystalline GaAs, a-GaAs displays topology and associated chemical disorder that have significant effects on device performance. This practical consideration has motivated numerous experimental studies by X-ray diffraction and EXAFS (extended X-ray absorption fine structure) on the structure of a-GaAs. These experiments provide invaluable structural information on both short and intermediate length scales. Despite extensive modeling efforts, understanding of structural correlations in a-GaAs is still inconclusive, in particular on intermediate length scales (> 4 Å).

Reliable interatomic potentials are essential ingredients of MD simulations. Interatomic potentials have been developed for semiconductors such as GaAs, AlAs, and InAs (Ebbsjö et al. 2000). The potentials were validated against experimental data and first-principles bulk electronic structure calculations on pressure-induced structural transformations, melting, and the cleavage energy. Predicted phonon dispersion of GaAs also agreed well with experimental data (Figure B3.1). MD results for the LO (longitudinal optical)-TO (transverse optical) phonon energy splitting at the Γ point (2.85 meV for GaAs and 2.46 meV for InAs) were also in reasonable agreement with experimental data (2.70 meV for GaAs and 2.46 meV for InAs).

Dr. Ingvar Ebbsjö of the Studsvik Neutron Research Laboratory, Sweden, has performed the best MD simulation of amorphous GaAs. This has been possible primarily because of the availability of an excellent interaction potential for GaAs. This interaction potential gives a very good description of the structural transformation under pressure for the bulk and nanoclusters.

Experimentally the intermediate range order (IRO), a manifestation of length scales (> 4 Å), shows itself as the first peak in X-ray and neutron static structure factors. For covalent glasses such as $GeSe_2$ and SiO_2, the IRO has been studied extensively by MD simulations (Campbell et al. 1999a). The first peak in these materials has been found to arise from atomic correlations at length scales in the range of 4–8 Å. The "universal" aspect of the first peak in these covalent glasses is represented by the product of the first-peak position, Q_1, and the nearest-neighbor atomic distance, r_1. For a large number of covalent glasses, $Q_1 r_1$ is around 2.5. Amorphous GaAs ($Q_1 r_1 \sim 5$) does not fit into this "universal" class. Nevertheless, a small value of Q_1 (~ 2 $Å^{-1}$) signifies that the first peak in a-GaAs represents correlations beyond near-neighbor distances.

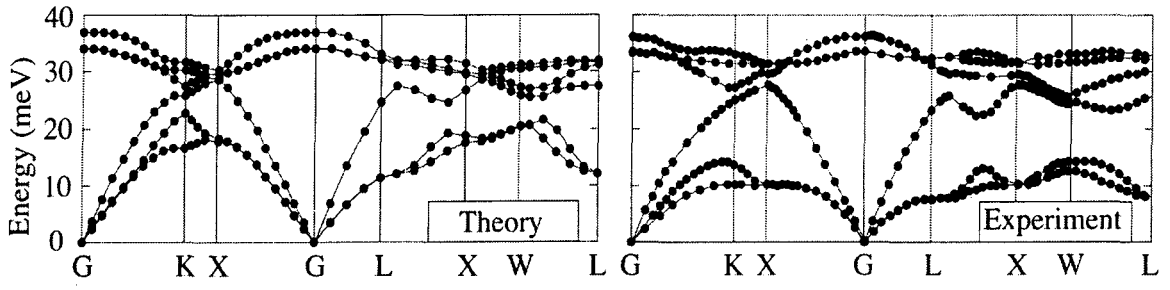


Fig. B3.1. Theoretical and experimental phonon dispersion of zinc-blende GaAs (Strauch and Dorner 1990).

Structural correlations in amorphous GaAs have been investigated with MD simulations using the new interatomic potential function (Ebbsjö et al. 2000). The calculated static structure factor of amorphous GaAs, in particular the height and width of the first peak which is a signature of the intermediate-range correlations, is in excellent agreement with X-ray diffraction experiments, as shown in Fig. B3.2. Atomistic topology on intermediate length scales is elucidated through the analyses of shortest-path rings, partial static structure factors, and bond-angle distributions. From the simulation data, the distribution of coordination numbers in a-GaAs, and the average coordination number, 3.92, are calculated and are found to be in close agreement with EXAFS results, 3.85. The calculated energy difference between crystalline and amorphous systems is also in good agreement with electronic-structure calculations.

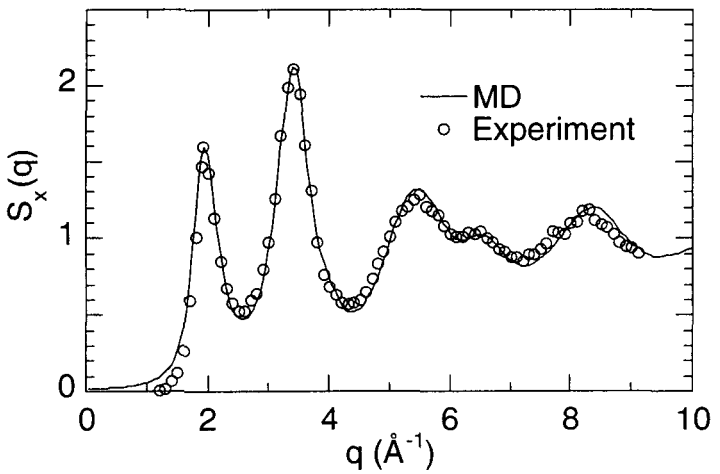


Fig. B3.2. X-ray static structure factor of amorphous GaAs at density 5.11 g/cm^3. The solid curve and circles represent MD and experimental data, respectively (Udron et al. 1991; Ebbsjö et al. 2000).

Oxidation of Aluminum Nanoclusters

The first successful MD simulation of oxidation dynamics of an Al nanoparticle was performed by Dr. Timothy Campbell (now at the Logicon, Inc., NAVO MSRC PET, Stennis Space Center, MS). The MD simulations (Campbell *et al.* 1999b) are based on a reliable interaction scheme developed at the Naval Research Laboratory by Streitz and Mintmire (1994), which can successfully describe a wide range of

physical properties of both metallic and ceramic systems. This scheme is capable of treating bond formation and bond breakage, and changes in charge transfer, as the atoms move and their local environments are altered. The MD simulations provide a detailed picture of the rapid evolution and culmination of the surface oxide thickness, local stresses, and atomic diffusivities (Fig. B3.3).

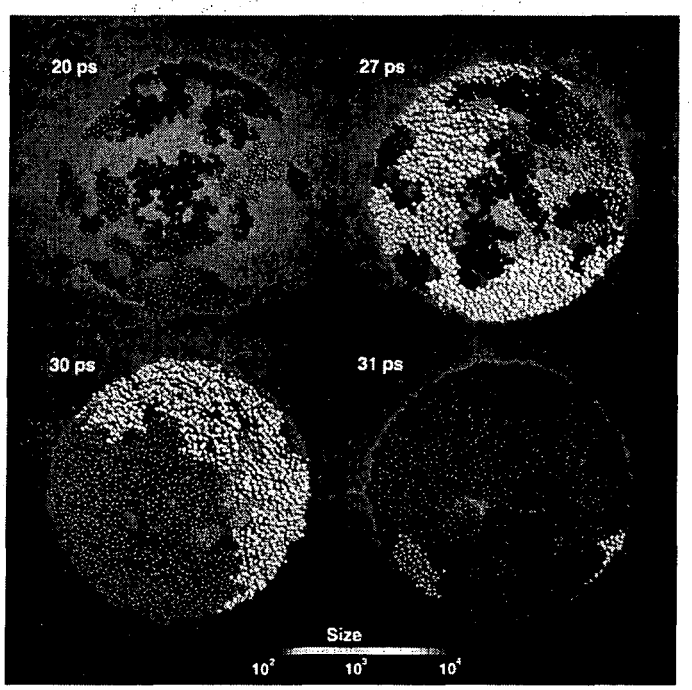


Fig. B3.3. Initial stages of oxidation of an Al nanoparticle. Size distributions of OAl_4 clusters between 20 and 31 ps are shown. Clearly, the clusters coalesce and percolate rapidly (Campbell *et al.* 1999b).

Sintering of Silicon Nitride Nanoclusters

Sintering is key to a number of advanced technologies. For example, multilayer ceramic integrated circuits (MCIC) are attracting much attention as an effective way to integrate discrete components for high-frequency wireless communication equipment (Wilcox *et al.* 1997). The major challenge in MCIC is to control constrained sintering of laminated ceramic multilayers to obtain mechanically stable products with desired properties. Computer simulations of MCIC is of particular interest to companies such as Motorola, Texas Instruments (TI), and other related companies in the area of wireless communication technologies.

The first simulations of sintering of ceramic nanoclusters have been carried out by Dr. Kenji Tsuruta of Okayama University, Japan. The MD simulations were performed to study sintering of Si_3N_4 nanoclusters, each cluster consisting of 20,335 atoms (Tsuruta et al. 1996). The simulations provided a microscopic view of anisotropic neck formation during early stages of sintering (Fig. B3.4).

MD simulations were also performed to study sintering between three nanoclusters. For nanocrystals, a significant rearrangement of the nanocrystals occurred within 100 ps, followed by the onset of neck formation. Amorphous nanoclusters aligned along a straight line were simulated. Within 100 ps, observed relative motion of the clusters was observed. In the next 100 ps, a symmetric neck formed between each pair of clusters and, thereafter, the relative motion subsided. The simulation showed a chain-like structure, which has been observed experimentally as well.

Dynamic Fracture and Stresses in Ceramics

Dr. Hideaki Kikuchi of the CCLMS carried out MD simulations of dynamic fracture in crystalline SiC at various temperatures using large-scale MD simulations on 256 nodes of a Cray T3E at DoD's Naval Oceanographic Office (NAVO). The simulations were performed with a reliable interatomic interaction

scheme that incorporates the ionic and covalent effects in the system. To test the validity of the potentials, he calculated lattice constants, cohesive energies, elastic moduli, melting temperatures, and phonon densities-of-states of the crystalline systems. MD results were in good agreement with experimental values.

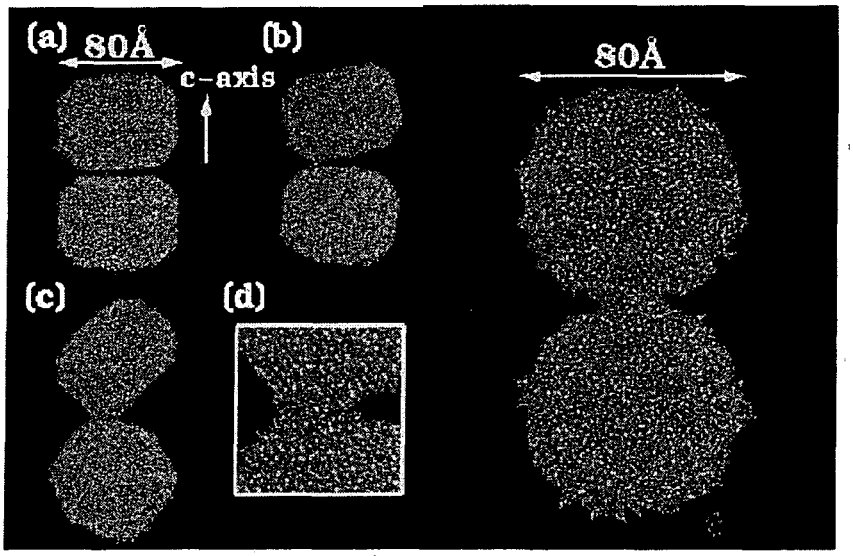


Fig. B3.4. (Left panel) Snapshots of Si_3N_4 nanocrystals at 2,000 K: (a) at time $t = 0$; (b) after 40 ps; (c) after 100 ps; and (d) close-up of the neck region after 200 ps. Spheres denote Si and N atoms, and lines represent Si-N bonds. (Right panel) Snapshot of sintered amorphous nanoclusters after 700 ps (Tsuruta et al. 1996).

Dynamic fracture simulations in SiC were performed on a thin-strip sample of 30 million atoms (Kodiyalem 2001). For this geometry, the mechanical energy release rate, G, of the system can be calculated from the knowledge of the applied strain, ε, and the value of the stress, σ, far ahead of the crack tip: $G = W\sigma\varepsilon/2$. In addition to the mechanical energy release rate, the crack-tip velocity and local stress distribution were monitored at various temperatures. The MD simulations revealed that large shear stresses close to the crack tip caused cleavage fracture at room temperature (Fig. B3.5). At elevated temperatures, dissipative effects due to dislocations, micropore formation and coalescence, and crack deflection caused stresses to spread out all over the system, thereby increasing the value of G.

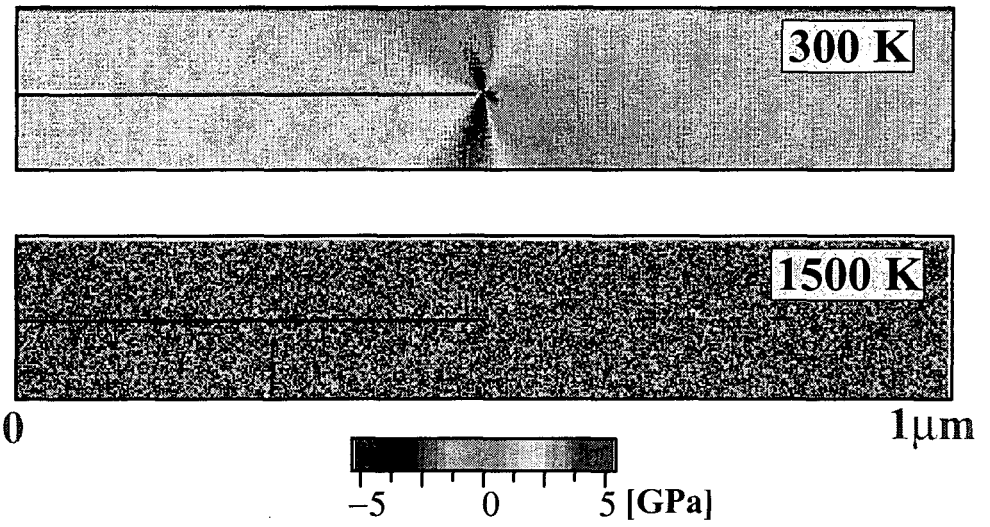


Fig. B3.5. Shear stress distributions in 30-million-atom MD simulations of SiC at 300 K and 1500 K.

Dynamic Fracture in Nanophase Si_3N_4

Understanding the role of microstructures in fracture is one of the most challenging problems in materials science. Nanophase materials are ideal systems for examining this issue at the atomistic level because microstructures in these materials are only a few nanometers in dimension. The morphological and dynamic aspects of fracture in nanophase Si_3N_4 have been investigated using 1.08 million particle MD simulations (Kalia et al. 1997b).

Snapshots of the crack front at various values of the external strain are shown in Fig. B3.6. The crack front consists of pores connected to the notch. Figure B3.6(a) shows a snapshot taken 10 ps after the notch was inserted. Initial development of the crack front and the growth of a few crack branches were observed in the system. These local branches and nanoclusters tended to arrest the motion of the crack front, and further crack propagation was only possible if additional strain was applied. Figure B3.6(b) shows a snapshot of the crack front in the nanophase system under 11% strain. Comparing with Fig. B3.6(a), it is evident that the crack front had advanced significantly and coalesced with the pores in the center. Also, pores and interfacial regions allowed the crack front to meander and form a branched structure. Figure B3.6(c) shows a crack-front snapshot 10 ps after the system was strained to 14%.

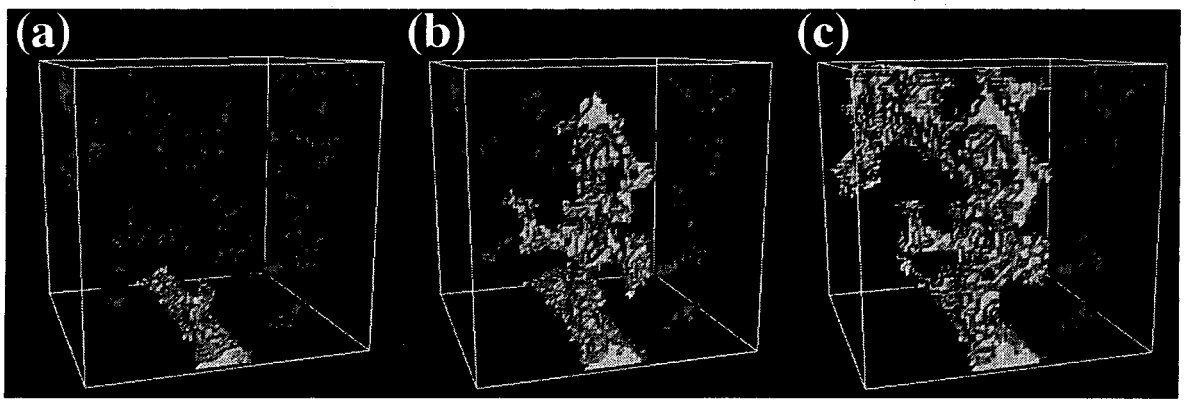


Fig. B3.6. Snapshots of the crack front (light gray) along with large (>6.4 nm^3) isolated pores (dark gray) in the strained nanophase system: (a) initial notch with crack branches and pores in the system under an applied strain of 5%; (b) the primary crack front after the strain is increased to 11%; and (c) the primary and secondary crack fronts at 14% strain on the system (Kalia et al. 1997b).

At that time, a secondary crack (top left-hand corner of the figure) with several local branches merged with the primary crack. With further increase in the strain, the secondary front advanced toward the initial notch while the crack kept growing laterally. When the strain reached 30%, the crack finally separated the material into two disconnected parts, and the system was thus completely fractured. The critical strain (30%) at which the nanophase system fractures was enormous (an order of magnitude) compared to what the crystalline Si_3N_4 system can sustain (3%). This was due to (1) multiple crack branching, (2) plastic deformation in interfacial regions, and (3) deflection and arrest of cracks by nanoclusters. None of these mechanisms were operative in the crystalline system. This demonstrates the dramatic effect of nanostructures on fracture.

Nanoindentation

Nanoindentation testing is a unique probe of mechanical properties of materials. Typically, an atomic force microscope tip is modified to indent the surface of a very thin film. The resulting damage is used to rank the ability of the material to withstand plastic damage against that of other materials. In addition, a load-displacement curve is constructed from the measured force at each displacement, and the elastic modulus in the direction of the indent can be measured from the initial part of the unloading curve. Commercial nanoindenting apparatus typically has a force resolution of ±75 nN and depth resolution of ±0.1 nm.

Walsh investigated nanoindentation of silicon nitride thin films using 10-million-atom parallel MD simulations (Walsh et al. 2000). Several simulations were performed to determine temperature effects, load-

rate effects, and simulation-size effects in crystalline and amorphous silicon nitride. The simulations were run on several different parallel platforms, including the 256- and 128-node IBM SPs at the U.S. Army Engineer Research and Development Center (ERDC), the 1,088-node Cray T3E at NAVO, and the 512-node Origin 2000 at Aeronautical Systems Center (ASC).

The films had lateral dimensions of 60.6 nm × 60.6 nm and thickness of 30 nm. Silicon nitride was simulated with an interatomic potential that has been thoroughly validated by comparison with experiment. The films were indented with a square-based pyramidal indenter to maximum depths of 8-9 nm. Figure B3.7 shows a comparison of load-displacement curves at three different temperatures. The hardness at 300 K (approximately 50 GPa) was reduced to approximately 90% of its value at 2000 K. Figure B3.8 shows the local pressure distribution directly under the indenter for the fully loaded and fully unloaded configurations. These pressure images were used in conjunction with local bond-angle calculations to characterize a process of local amorphization under the indenter, which is arrested by either piling up of material along the indenter edges or by cracking under the indenter corners.

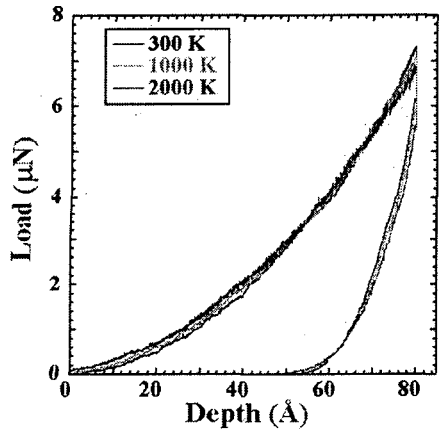


Fig. B3.7. Load-displacement curves for 300 K, 1000 K, and 2000 K nanoindentation simulations.

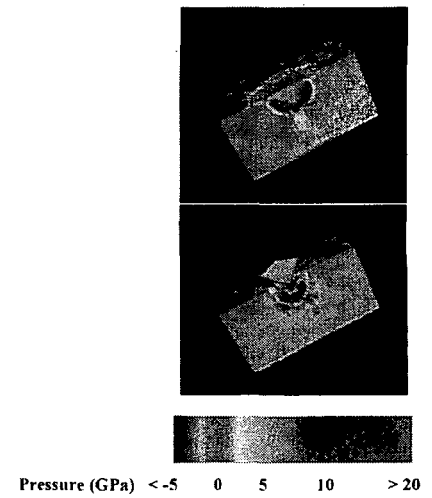


Fig. B3.8. Local pressure distributions for (top) fully loaded and (bottom) unloaded configurations.

Ceramic Fiber Composites

Physical properties of composite materials often exhibit synergistic enhancement. For example, the fracture toughness of a fiber composite is much larger than a linear combination of the toughness values of the constituent materials. This enhanced toughness has been attributed to the frictional work associated with pulling out of fibers, which suggests that tough composites can be designed by combining strong fibers with weak fiber-matrix interfaces.

MD simulations (Fig. B3.9) were performed by Dr. Aiichiro Nakano of the CCLMS to investigate the atomistic toughening mechanisms in Si_3N_4 ceramic matrix (bulk modulus 285 GPa) reinforced with SiC fibers (bulk modulus 220 GPa, 16 vol. % fibers) coated with amorphous silica (bulk modulus 36 GPa) (Vashishta et al. 1999). The simulations involving 10 million atoms were performed on 128 SGI Origin 2000 processors. Fiber-reinforcement was found to increase the fracture toughness by a factor of two. The atomic-stress distribution showed an enhancement of shear stresses at the interfaces. The enhanced toughness resulted from frictional work during the pullout of the fibers. Immersive visualization of these simulations revealed a rich diversity of atomistic processes including fiber rupture and emission of molecular fragments, which must be taken into account in the design of tough ceramic composites.

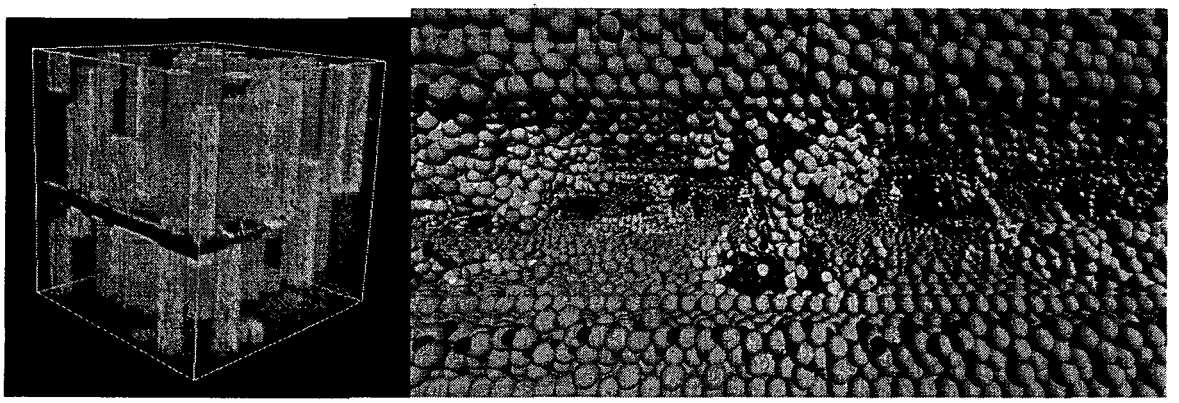


Fig. B3.9. (Left panel) Fractured silicon nitride ceramic reinforced with silica-coated silicon carbide fibers. (Right panel) Close-up of the fractured composite system. Small spheres represent silicon atoms and large spheres represent nitrogen, carbon, and oxygen atoms (Vashishta et al. 1999).

REMARKS

The examples represent an academic grand-challenge approach to modeling hard condensed matter, seeking out the largest computational resources and using them to attack a limited number of very big problems rather than many small ones. The liaison between academic departments of physics and of computer science is a noteworthy part of that approach.

REFERENCES

Campbell, T.J., R.K. Kalia, A. Nakano, F. Shimojo, S. Ogata, and P. Vashishta. 1999a. Structural correlations and mechanical behavior in nanophase silica glasses. *Phys. Rev. Lett.* 82, 4018.

Campbell, T.J., R.K. Kalia, A. Nakano, P. Vashishta, S. Ogata, and S. Rodgers. 1999b. Dynamics of oxidation of aluminum nanoclusters using variable charge molecular-dynamics simulations on parallel computers. *Phys. Rev. Lett.* 82, 4866.

Ebbsjö, I., R.K. Kalia, A. Nakano, J.P. Rino, and P. Vashishta. 2000. Topology of amorphous gallium arsenide on intermediate length scales: a molecular dynamics study. *J. Appl. Phys.* 87, 7708.

Kalia, R.K., A. Nakano, A. Omeltchenko, K. Tsuruta, and P. Vashishta. 1997b. Role of ultrafine microstructures in dynamic fracture in nanophase silicon nitride. *Phys. Rev. Lett.* 78, 2144.

Kodiyalam, S., R.K. Kalia, H. Kikuchi, A. Nakano, F. Shimojo and P. Vashishta. 2001. Pressure induced structural transformation in gallium arsenide nanocrystals," *Phys. Rev. Lett.*, to be published.

Strauch, D., and B. Dorner. 1990. *J. Phys.: Condens. Matter* 2, 1457.

Streitz, F.H., and J.W. Mintmire. 1994. Electrostatic potentials for metal-oxide surfaces and interfaces. *Phys. Rev. B* 50, 11996.

Tsuruta, K., A. Omeltchenko, R.K. Kalia, and P. Vashishta. 1996. Early stages of sintering of silicon nitride nanoclusters: A molecular dynamics study on parallel machines. *Europhys. Lett.* 33, 441.

Vashishta, P., R.K. Kalia, and A. Nakano. 1999. Large-scale atomistic simulation of dynamic fracture. *Comput. Sci. Eng.* 1 (5), 56.

Walsh, P., R.K. Kalia, A. Nakano, P. Vashishta, and S. Saini. 2000. Amorphization and anisotropic fracture dynamics during nanoindentation of silicon nitride—a multi-million atom molecular dynamics study. *Appl. Phys. Lett.*, in press.

Site: **Lucent Technology/Bell Laboratories**
700 Mountain Avenue
Murray Hill, NJ 07974-0636
http://www.lucent.com/

Date Visited: 17 January 2000

WTEC Attendees: P. Vashishta (report author), P.A. Kollman, K. Morokuma, E.B. Stechel, P.R. Westmoreland

Hosts:
Dr. Conor Rafferty (Coordinating host), Head of Silicon Processing Research Department; (908) 582-3575, FAX (908) 582-4228; conor@lucent.com (conor@agere.com)
Dr. Frieder Baumann, Silicon Processing Research; (908) 582-2191, FAX (908) 582-4228; frieder@lucent.com
Dr. George H. Gilmer, Silicon Processing Research; (908) 582-5547, FAX (908) 582-4228, ghg@lucent.com
Dr. Mark S. Hybertsen, Semiconductor Photonic Components; (908) 582-3628, FAX (908) 582-7660; msh@lucent.com
Dr. Krishnan Raghavachari, Photonics Materials; (908) 582-6783, FAX (908) 582-3958, krish@lucent.com
Dr. Frank Stillinger, Materials Chemistry Research; (908) 582-4500, FAX (908) 582-3958, fhs@bell-labs.com

BACKGROUND

Lucent was founded in 1996 as a spin-off from AT&T. The company's slogan is "We make the things that make communications work." The company has headquarters in Murray Hill, NJ, in the same building as the research laboratory. Lucent is world's largest manufacturer of communication equipment—local, long distance, international and satellite. Lucent had FY 1999 revenues of $38.3 billion and net income of $3.8 billion. 11% of the revenues are spent on R&D and there is an understanding that 1% of revenues will be spent on basic research. Lucent has 153,000 employees world wide—77% in the United States. The company has had an average growth rate of 25%.

Lucent-Bell Laboratories is the one of the premier industrial research laboratories, not only in the United States but also in the world. AT&T Bell Laboratories, the predecessor of Lucent-Bell Laboratories, was established in 1925 to carry out research in electronic devices for telephone communications. In the 1940s, it moved from New York to its current location in Murray Hill, a small town in New Jersey, some 30 miles west of New York City.

Lucent Research has 30,000 staff in Research and Development, out of which 3,000 are in research. There are several locations worldwide. Largest are at Murray Hill and Holmdel in New Jersey and a new location in California, BLSV (Silicon Valley). The laboratory has four divisions—physical sciences, computing and mathematical sciences, communications sciences, and packet switching (BLSV). The laboratory has 25,000 patents and produces more than 1,000 patents per year.

Based on the research carried out at the laboratory, six Nobel Prizes in Physics have been awarded to its scientists. Major discoveries/inventions of the laboratory include the transistor, solar cell, CCD, radio astronomy, laser, LED, transparent fiber, information theory, molecular beam epitaxy, the Big Bang, Unix, C, and C++. The best known invention of Bell Labs is the transistor. Bardeen and Brattain discovered the point-contact transistor in December 1945 at the Murray Hill location. Bardeen, Brattain, and Shockley were awarded the Nobel Prize for their discovery in 1956, the second Nobel for Bell Labs. In 1945, no one knew how much impact the transistor would have on business, education and culture.

The motivation of Bell Labs still is to carry out valuable basic research in an industrial framework. Primary research areas of the Lucent-Bell Laboratories, established in 1996 after the spin-off from its parent

company, AT&T, are related to core technology businesses of Lucent—electronic devices and communication products. Areas of research include semiconductor physics, lasers, optoelectronics, and hardware and software development for fiber and wireless communication.

In the Physical Science Division, there are 575 staff located centrally at the Murray Hill location. In the late 1980s the peak size of staff was about 700. In current years the research emphasis has been shifting to software/networking, reflecting the core technologies of Lucent. The effort is divided into four laboratories—physical research laboratory ("fundamental" lab), silicon electronics research lab, photonics research lab, and materials processing research lab. Its current research directions include:

- Optical transmission: toward 40 terabit, free-space transmission
- Optical switching: micromirrors from lab to product in 15 months
- Microelectronics: SCALPEL electron beam lithography; vertical transistor, 100 MHz DSPs
- Electronic paper: plastic transistors, electronic ink
- Wireless: geolocation, data networking, multipath
- Networking: protocols for 10 gigabit ethernet, firewalls
- Signal processing: speaker identification, video watermarking
- Mathematics: quantum computing

In July 2000, Lucent announced that it intended to spin off its Microelectronics Group by the summer of 2001. This new company will be named Agere Systems. This spin-off will essentially split Bell Labs Physical Sciences Research down the middle. Of about 500+ physical science researchers, 275 are leaving with Agere. Through all the previous splits of the historical AT&T Bell Labs (divestiture in 1983, Lucent spin-off in 1995), the physical sciences area remained mostly intact. Generally, personnel associated with lasers, transistors, and electronic circuits will go with Agere, while the remainder will stay with Lucent, including those working on projects such as plastic transistors, bio-computing, polymer chemistry and some MEMS.

R&D ACTIVITIES IN MOLECULAR AND MATERIALS MODELING

After an overview presentation on Lucent's corporate and research structure by Dr. Rafferty, six research presentations (Drs. Rafferty, Gilmer, Baumann, Hybertsen, Raghavachari, and Stillinger) were made to the panel members. These scientists came from diverse backgrounds—electrical engineering, solid state physics, physical chemistry, and quantum chemistry. Common features of nearly all the modeling and simulation research discussed were (1) that the project involved both simulation and experimental scientists and (2) there was some contact with relevant processing or technology groups within the company.

Dr. Conor Rafferty described research on deposition of barrier layer films—models and experimental validation. This effort involved deposition in vias and trenches to examine the nature of resulting coverage, porosity on sidewalls, connectivity of pores, sidewall and bottom coverage and nature of diffusion and its effects. ADEPT Monte Carlo model simulations were carried out on single DEC Alpha workstations. Simulation results were compared with X-ray reflectance experiments, atomic force microscope analysis of films, and high-resolution transmission electron micrographs.

Dr. George Gilmer discussed atomistic modeling of ion implantation and annealing of silicon devices. This project involved continuum modeling using PROPHET by Drs. Rafferty and H. H. Vuong, Monte Carlo simulations by Dr. Gilmer and Dr. L. Pelaz and M. Jaraiz (University of Valladolid), molecular dynamics by Dr. Gilmer and Dr. T. Diaz de la Rubia (Lawrence Livermore National Laboratory), and experiments by Dr. H.-J. Gossmann. Ion and cluster implantation, subsequent annealing, and amorphization in silicon were investigated using molecular dynamics; interstitial clustering in silicon and the nature of {311} defects and effect of ion dose were investigated using Monte Carlo methods. Dopant diffusion and clustering (deactivation) were also investigated.

Dr. Frieder Baumann discussed a very interesting project which is a part of jointly funded NSF and DARPA program called "Virtual Integrated Prototyping (VIP)". Several multi-institutional teams are funded in the VIP program. Lucent is the industrial partner in the project team lead by Professor Joe Green (materials scientist, University of Illinois, Urbana). The goal of the VIP program is to bring together mathematicians and materials scientists for collaborative research on materials problems of industrial interest. The Lucent team involved Drs. Baumann and Gilmer, and Drs. Jacques Dalla Torre and Peter O'Sullivan. To carry out multiscale modeling of barrier layer deposition, the team combined molecular dynamics using embedded atom methods, ADEPT hard sphere simulations using Monte Carlo, and continuum models using level set methods. The output from one level of simulation was used as input into the next level. Examples of problems studied include: (1) texture dependence of growing film (columnar growth versus porous films), (2) effects of inclined substrate (important at side walls of features), and (3) effects of imperfect wetting of underlying substrate (influence of substrate on film texture).

Dr. Mark Hybertsen discussed his modeling and simulation effort which was almost entirely directed toward semiconductor laser technology—a very important and profitable product area for Lucent. He described a quantum molecular dynamics simulation of Si-SiO_2 interface carried out in collaboration with Parrinello and Car (Switzerland) in 1998. This exploratory simulation, even though quite small in size, provided considerable insight into the nature of this very important interface. The majority of research described by Dr. Hybertsen was in the area of microscopic laser device simulation (quantum optics, carrier transport, carrier recombination, quantum well capture and quantum well optical properties) and impact of dopant diffusion in the materials involved in the active region of the laser. Future directions will involve alloy properties of novel materials such as InGaAsN for semiconductor optoelectronic technology.

Dr. Krishnan Raghavachari is an expert in highly accurate quantum chemical calculations. He is using a variety of techniques depending on the size of the simulation and the accuracy needed to make a meaningful comparison with experimental results to validate the model. Quantum chemistry capabilities include: (1) accurate geometries (bond lengths to 1-2%, angles to within a few degrees), (2) vibrational frequencies (within 5–10%), (3) trends in vibrational frequencies (within a few wavenumbers), (4) reaction energies, activation barriers (within 0.1–0.2 eV), and (5) electronic excitation energies of small systems (within 0.1–0.2 eV). Cluster approaches are used to study bulk, surface, and interface phenomena. To study silicon, silica glass, and silicon/silica interfaces, a suitable local region is chosen and the dangling bonds at the outer region of the cluster are terminated with hydrogen. Questions the theoretical model must satisfy before reliable answers can be extracted are: (a) is the cluster large enough to attain convergence, (b) is there sufficient distribution of structures when studying defects in glasses, and (c) since convergence for excited states is difficult for large system sizes, is the system sufficiently large to extract meaningful results? This approach has been applied successfully to vibrational problems (vibrations above 600 cm^{-1} to assign structures, hydrogen vibrations on Si(100) surface, and oxidation of Si(100) surface). Mechanistic problems such as HF etching of silicon surfaces have also been studied.

Dr. Frank Stillinger is one of the best known and senior members of the staff at Lucent-Bell Laboratories. He prefaced his remarks by pointing out that he chaired a study entitled "Mathematical Research in Materials Science" carried out in 1993 under the auspices on the National Research Council. He is also interacting with the Princeton Materials Research Center and spends one day a week at Princeton University consulting with the Center.

Dr. Stillinger described his research interests as a reflection of the history of modeling and simulation at Bell Laboratories since 1967. Key features and studies have included the following:

1. Water (1967)
2. Fluorinated hydrocarbon DNA analogs (1978)
3. Aqueous surfactants (1980)
4. Phosphoric acids (1982)
5. $Ni_{80}P_{20}$ amorphous alloys (1982)
6. Silicon crystal and liquid (1985)
7. Sulfur polymerization (1986)
8. F_2 into F + F (1989)
9. Glycerol (1989)

10. Silicon etching by F (1989)
11. Organic dyes in glycerol (1990)
12. Liquid Si_2F_6 (1991)
13. Cyclohexane (1991)
14. Poly-L-Alanine (1992)
15. Crystalline ^4He (1995)
16. Silicic acid condensation (1996)
17. Ortho-terphenyl (1998)
18. Pentane isomers (1999)
19. SiO_2 sol-gel processes (2000)

DISCUSSION AND ADDITIONAL COMMENTS

- Before the formation of Lucent, about 10% of the 700 researchers at Bell Laboratories were theorists. Now there seems to be some reduction in numbers due to attrition. Most of the increase in research staff in recent years has been in the area of communication and software.
- The measure of success has changed. Before Lucent was spun off, excellence in research for its own sake was sufficient. Now, research which helps solve a technological issue for the company is most valued.
- Theory Department (Head, Don Hamann) is the only pure theory department left at Lucent.
- Simulation effort at Lucent is integrated with experimental and processing programs. Simulation effort is neither decreasing nor increasing. It is in steady state.
- Hardware: distributed computing resources are managed by groups. The theory group headed by Dr. Hamann has a 64-CPU PC cluster. Dr. Raghavachari has an 8-processor system from SGI.
- Software: For quantum chemistry, Gaussian is used most, although there is some development of quantum chemistry techniques and codes. Polymer modelers have some commercial software, but many codes used in simulations are developed at Lucent.
- Validation and verification: Ultimately, the simulation software should be able to run on a PC if the simulation is developed for a processing environment.
- Research atmosphere and interactions: There are no cultural barriers between experimental and theoretical researchers. There is excellent collaboration between the two at Lucent.

Site: **Marathon Oil Company**
7400 South Broadway
Littleton, Colorado 80122
303-347-5565, FAX 303-347-5551
http://www.marathon.com/

Interview Dates: 9 February 2000, 9 October 2000

Interviewer: P.R. Westmoreland

Contacts: Dr. Mark A. Plummer, plummer@columbine.net
Mr. John L. Jechura, john_jechura@nrel.gov

BACKGROUND

Marathon Oil Company is part of USX Corporation (Pittsburgh, PA), which also operates U.S. Steel, the nation's largest producer of steel. Marathon was founded in 1887 as The Ohio Oil Company, bought by Standard Oil in 1889, divested as independent in the 1911 break-up of the Standard Oil Trust, and renamed Marathon in 1962. In 1982 it became a wholly owned subsidiary of U.S. Steel, which subsequently became USX.

Marathon currently has about 4000 employees, excluding foreign and subsidiary company employees. It is engaged in exploration, production (in top ten of U.S.-based oil producers), natural gas (1 billion cubic feet per day, much shipped to Japan as LNG), refining (seven refineries constituting 6% of U.S. refining capacity, 935,000 barrels/day), transportation (5,000 miles of oil pipelines), and marketing (Marathon, Ashland, Speedway, Starvin' Marvin brands).

MOLECULAR MODELING R&D ACTIVITIES

Mark Plummer has been responsible for computational chemistry for catalyst applications, while John Jechura has worked to apply molecular models for process optimization as developed by Michael Klein of Rutgers University (formerly Delaware). Earlier, this activity had been using computational chemistry to study hydrogenation catalysis (starting about 1995), mainly using Hyperchem (Hypercube, Inc.). Marathon researchers have extended this work to liquid-liquid extraction for removing sulfur from gasoline, but they have not started using molecular models for process optimization.

Marathon's molecular modeling was carried out at its Denver research center, which Marathon decided to close down, effective December 2000.

Site: **Merck & Co., Inc.**
Merck Research Laboratories, RY50-SW1DD
P.O. Box 2000, 126 E. Lincoln Ave.
Rahway, NJ 07065
Tel: (732) 594-4000
http://www.merck.com/

Interview Dates: 18 January 2000 (data reflects state in December 2000)

Interviewer: P.A. Kollman (report author), R. Miranda, K. Morokuma, E.B. Stechel, P. Vashista, P.R. Westmoreland

Contacts: Simon Kearsley (principal contact), Head, Molecular Systems Group, 732-594-3656, FAX 732-594-0670, skk@merck.com
Laurie Castonguay, Application Group Manager (Rahway), 732-594-6301, FAX 732-594-4224, castong@merck.com
J. Chris Culberson, Application Group Manager (West Point), 215-652-7537, FAX 215-652-4625, jcc@merck.com
Brad Feuston, Development Group, 215-652-5048, FAX 215-652-4625, bradley_feuston@merck.com
Eugene Fluder, HPC Manager, 732-594-5074, FAX 732-594-4224, fluder@merck.com

BACKGROUND

Merck is one of the leading, and one of the largest, pharmaceutical companies in the United States and in the world. Sales in 1999 were $33 billion, and the company employed 62,300 people. From 1995 through 1999, Merck launched 15 new medicines.

In the United States, Merck has two major research sites: Rahway, New Jersey, and West Point, Pennsylvania. Abroad, it has smaller sites in Terlings Park, UK; Montreal, Merck Frosst, Canada; and IRBM, Italy.

MOLECULAR MODELING R&D ACTIVITIES

The modeling group in New Jersey and Pennsylvania, also called Molecular Systems, currently numbers 27 people. Roughly half are developing new methods and half are applying modeling to drug discovery projects. A portion of the development group is involved in computer maintenance. Kearsley is the head of the whole group and reports to a vice president of chemistry at Rahway. Castonguay and Culberson are managers of the applications and methods groups in Rahway and West Point, respectively. Fluder has responsibilities for high performance computing at all sites and aspects of methods development.

On average each person in the modeling group works on one or two major projects and a number of minor ones. A major project is usually one that has a laboratory objective of producing a Safety Assessment Candidate (an SAC is a compound that is deemed worthy of preclinical safety analysis and is likely to go on to be a drug). Application modelers are expected to make tangible contributions or suggestions that will influence the project in a one- to two-week time frame. Computations that take longer than this will often miss their window of impact. If the modeler is primarily in methods development, a major project may well involve writing a substantial piece of software.

For the past decade and for the foreseeable future, modelers make a substantial impact on projects by finding new leads and creating Structure Activity Relationships (SAR) through chemical database mining. The success of their chemical database mining contributions depended on the methods group creating robust or improved software for efficient virtual screening of many compounds (e.g., topological similarity searches,

molecular superposition algorithms and docking methods). This requires development of efficient, relevant scoring functions and considerable computer resources.

It is felt that the modeling group has a relatively unique niche among the project scientists. As a group, it is exposed to many projects and therefore more able to see relationships between projects. Many drug discovery programs are jump-started from related programs. Thus, one of the responsibilities and great challenges for this group is to formalize this cross-fertilization of knowledge between drug discovery projects.

Merck has always had a major effort in methods development. They believe this gives them a competitive edge. They prefer not to rely on commercial or academic software.

DISCUSSION

Percentage of Targets with Known Structures

About 25% of the targets are known or can be accurately built by homology. That percentage hasn't changed recently, but the variety of targets has, in that there are many more X-ray structures available. Merck has an active effort in modeling G-protein-coupled receptor systems. Qualitative results using data from molecular biology (mutating the receptors) and pharmacology (considering different ligands as agonists or antagonists) on such systems are valued as much as the more quantitative modeling expected from crystal structure data.

Relationship to Other Groups

The modeling effort at Merck goes back to the 1970s. At times, chemists were encouraged and educated to do their own modeling. However, the current approach is for modeling to be done by the member of the project team who is from Molecular Systems. Chemists and biologists are not encouraged or discouraged from doing their own molecular modeling, and can if they want to; but it is considered to be more efficient for them to concentrate on what they do best and work together with a modeler to decide what direction the project should take.

Currently the modeling group provides support to 450 chemists and around twice as many biologists (molecular biology, metabolism and pharmacology). Modeling contributions are generally sought out by chemists and biologists. The role of the modeler is very clear when it comes to chemical database mining, assisting decision making when there is a lack of experimental data, and providing an alternative view of the SAR. Modelers are regarded as scientists and an integral part of the project team.

Credit and How Success is Evaluated

The focus in basic research is to come up with an SAC, and the team that does this (which usually includes a modeler) is recognized accordingly. As always, the challenge is to evaluate and reward team contributions, and important contributions that are more tangential to producing an SAC. Publication is encouraged and required for promotion.

Software/Hardware Capabilities and Needs

Merck uses both external and internal software. Merck researchers have a long history of creating their own modeling systems; they are in their third major revision. The current one, MIX (Methods Development Workbench, or Modeling In UNIX), focuses on facilitating methods development, and strives to complement the available software from the ISVs (independent software vendors). The developers are responsible for maintaining and evolving MIX. This maintenance overhead is more than compensated for by custom methods development, timely enhancements and bug fixes, and general interoperability benefits. Still, they use software from the ISVs because they want to make sure they have all the capabilities they might need (SYBYL, QUANTA/CHARMM, Gaussian 98, Schroedinger's Macromodel and some others).

For CPU servers they have 36 processors of Cray SV1 and an 8-processor T90, as well as a desktop SGIs for everyone. At the UK and Canadian sites, they have SGI Origin 2000s with a total of 48 processors. For file and security services, they use a DFS/DCE environment, and use an ADSM robotic backup/archival system. These HPC resources are made available to anyone in research.

Challenges Currently and in the Future

Modeling as an applied discipline within the pharmaceutical arena must transcend traditional computational chemistry, and embrace more qualitative and data driven approaches. At the same time it must balance this with advances in theory, so that scoring functions will improve. Treating vast volumes of data remains a challenge. Further improvement in modeling ADME (absorption, distribution, metabolism and excretion) is extremely important. It was felt that the biggest challenge, considering the number of research directions one could take, is to focus on advances that will make a significant impact on drug discovery in a timely manner.

Site: **Molecular Simulations, Inc. (MSI; now Accelrys)**
Corporate Headquarters
9685 Scranton Road
San Diego, CA 92121-3752
(858) 458-9990, (858) 458-0136 FAX
http://www.msi.com/

Date Visited: December 6, 1999

Interviewer: A.M. Chaka (report author)

Hosts: Bruce E. Eichinger, Corporate Fellow, 858-799-5540, bruce@msi.com

BACKGROUND

Molecular Simulations Inc., a subsidiary of Pharmacopeia, Inc. (http://www.pcop.com; see separate report), is a leading provider of molecular modeling, simulation, and informatics software and services for both life and materials science research. MSI employs over 325 employees worldwide, with 190 at corporate headquarters in San Diego. Additional R&D facilities are located in Cambridge, UK. Approximately 150 employees hold PhDs.

The company has evolved by growth and acquisition. Early companies that later became part of MSI (Richon 1996 and updates) have included the following:

- Hare Research, founded by Dennis Hare in 1982 in Washington
- BioDesign Ltd., founded in 1984 by Barry Olafson, Stephen Mayo, and William Goddard of Caltech to sell the program Biograf
- Polygen, founded in Massachusetts in 1984 by Frank Momany, Jeffrey Wales, Jean-Loup Fayolle, and Andy Ferrara to sell the program QUANTA / CHARMm originally developed by the group of Martin Karplus at Harvard (Brooks et al. 1983)
- Biosym Technologies, founded in California in 1984 by Arnold T. Hagler and Donald McKay to sell Insight/Discover; its first consortium (1986) was focused on potential energy functions for molecular simulations
- Cambridge Molecular Design, founded in the United Kingdom in 1989 by Patrick Coulter to sell the software set Cerius
- BioCAD Corporation, founded in 1989 in California by Terry Smith, Michael Jacobs, and Steve Teig to sell Catalyst software

In 1991, BioDesign changed its name to Molecular Simulations, Inc., acquired Cambridge Molecular Design and its Cerius package, and then merged with Polygen; MSI acquired BioCad in 1994. In parallel, Biosym developed its Discover package, acquiring Hare Research and being acquired by Corning in 1992. Both Cerius and Discover included a code for visualization, basic construction of molecules, and coupling modules for molecular simulations and computational quantum chemistry. Corning merged MSI with Biosym in 1995 and sold the combined MSI to Pharmacopeia in 1998. In 2000, Pharmacopeia acquired the chemoinformatics company Synopsys Scientific Systems Ltd. (Leeds, UK) and the bioinformatics and cheminformatics software parts of Oxford Molecular Group Ltd. , which were to become part of MSI Life Sciences. As of June 1, 2001, Pharmacopeia's software subsidiaries were combined and named Accelrys.

Consortia of companies and academic institutions are still important parts of the company. Consortium members help guide software development directions, receive existing and advance releases of software, and participate in regular meetings with each other and MSI. There are four consortia in materials science: polymers (founded in 1989 and presently reporting 30 members), catalysis (1991, 28 members), formulations (1999, involving Unilever, Henkel, and Clorox as of August 2000), and molecular crystallization (2000, formed as a new phase of the 1997-2000 Pharmaceuticals Development Consortium, which had 21

members). The three consortia in life science cover combinatorial chemistry, functional genomics, and high-throughput crystallography.

MOLECULAR MODELING SOFTWARE

MSI conducts research in-house and also partners with academic and industrial collaborators. The company maintains a wide spectrum of software products for quantum chemistry, molecular simulations, and information technology, covering the following areas:

- Catalysis and sorption
- Chemical reactions
- Crystallization
- Formulations using QSAR
- Mesoscale modeling
- Pharmaceutical development
- Polymer science
- Quantum methods
- Combinatorial chemistry
- Rational drug design
- Protein bioinformatics
- Protein design and simulations
- Structural biology
- NMR spectroscopy
- Macromolecular X-ray crystallography
- 3D structural databases
- Visualization & communication

MSI software is currently installed at more than 3,500 commercial, academic, and government R&D sites in such industries as agrochemicals, automotive, biotechnology, chemicals, food, personal care products, petroleum, pharmaceuticals, plastics and rubber, and paints and coatings. Presently, it classifies its products as being for life science or materials science. Nonbiological chemical applications are included in the latter. A current thrust is to port many of its applications to PCs, where previously they have been available only on UNIX workstations and large computers.

As examples of their products' use and of the methods' use, they have developed a quarterly magazine (*Solutions*) and an extensive Web site (http://www.msi.com/solutions/cases/). They have also maintained user groups (e.g., European Materials Science Users Group, http://www.msi.com/user/groups/paris98/) and the consortia that were the foundation of Biosym.

DISCUSSION

The following is a report based on WTEC's interview with Dr. Bruce Eichinger, who currently holds the position of Corporate Fellow at MSI. In 1989 he came from the University of Washington to be the first director of Biosym's polymer consortium.

In the interview, he offered a wide range of his own assessments and comments on MSI's experiences and on industrial experiences with molecular modeling. These are his comments as an MSI employee, but they do not necessarily indicate MSI policy or official opinion. They also reflect more the chemicals and materials side of MSI, while the life sciences part of the company is discussed in the Pharmacopeia report. The following are his comments:

On Successful Application

> Industrial success and satisfaction have been mixed. Several companies that started back in the late 1980s have continued to build up their modeling expertise. Other companies have dabbled in modeling, have been disappointed, and have either contracted it out or abolished it altogether. The real issue has to do with the quality of people doing the modeling and the appropriate selection of problem areas, which again speaks to the quality of the people and their ability to pick problems. Working closely with experimentalists is absolutely essential.

One illusion is that molecular modeling generates results and products in a *de novo* sense. In the relatively early days of application, researchers and managers alike can fall prey to this easy but mistaken assumption. Usually, chemical problems don't get solved by just computing something. One needs a clearly defined strategy based on clear intuition. You have to solve 90% of the problem in your head and then do the calculation to test various hypotheses.

One of the things that computations provide is that approaching a problem with a computer requires a well-thought-out strategy. In the process of developing that strategy, however, you learn techniques that can be applied to experimental projects as well. Does modeling answer all the questions every step of the way? No. A lot of experimentalists don't do that either—they start collecting data. Data is well regarded and seen as an end in itself, and is often not evaluated in terms of how useful is it. I recall that back in 1990 a prominent polymer scientist stated that "You can get a lot of wrong answers with computations." Yes, but the same is true for experiments. It is essential to have a firm grasp of what you can and cannot do.

In some of the jargon we [modelers] use, getting alignment with the customer is essential for success. The modeler and customer must both agree on what is to be done—i.e., between what the modeler can accomplish and what the expectations are. In many cases of modeling, that alignment has been deliberately or unfortunately overlooked.

As for specific examples, companies often don't tell MSI what their key successes have been. MSI is not allowed to discuss contract work.

On Important Technologies for Modeling in Near and Long Term

Complex Reactions

Most of the new developments need to address real chemistry. That is, for chemical reactions we need to predict rates in situations where the chemistry is very complex, and it is difficult to sort through all the possibilities.

Transition States

I have a sneaking suspicion we really need a lot of work for transition states—not only in finding them, but in getting the energies right. This is a fundamental problem. A G2 data set does not exist for transition states. We need a good data set and many, many examples. It's just not there and is absolutely critical.

Long-Term Stability with Respect to Slow Processes

The other area is also related to chemicals, but it is how to predict the long-term stability of materials in general. I can't tell you how often we get asked to predict the lifetime of plastic X. How do we treat slow processes like oxidation, UV stabilization, physical aging, coupled together with fracture mechanics? This is a huge problem. It's too hard for most academics and hence doesn't get treated.

Polymer Properties

We are very close to being able to compute properties from first principles for pure polymers. But we still don't know what to do about composites and more complex systems. There exists an enormous need to build the representative structures for composites and the technology for handling the mechanics for those problems.

Fracture Mechanics

The area of fracture mechanics is a good example of the challenges facing predictions for composite materials. About 5 years ago we were inundated by pictures of pretty surfaces of fracturing Lennard-Jonesium (hard spheres interacting with a Lennard-Jones 6-12 potential). What ever happened to that? There was no technology transfer on that problem. Industrial problem solving will not happen when a teraflop machine is required. What is the role of these enormous calculations? If the goal is to provide support, insight, or validation, ok—it can be addressed by a few people in the world. But this type of massive computation will not have a role in practical industrial problem solving.

Defect Structures in Materials

Whether grain boundaries or dirt in polymers, treating defects is absolutely critical to achieving understanding in fracture mechanics. How do we come to grips with this when they [fractures] are, in essence, rare events? They can occur in 1 part per billion, but totally change the properties of the material. And how does one even begin to describe a dirt particle?

Combining Length Scales

The other big problem is how to combine length scales. There are a vast [number of] unanswered questions in this area.

On Issues for Government and Academia

With respect to making progress in these areas in academia, these problems are large and will require team efforts. Most of the work done in chemistry, however, is done by rugged individuals who are not used to working in teams. In addition, academic research follows the line of least resistance. If a problem is too hard, you think about a different one and get the student out. It must be understood that new sets of rules need to be applied to these very difficult problems. New guidelines of funding and expectations need to be set up. No one likes to claim success with negative results, but that is a critical element in science. A negative result is as good a learning experience as a positive one, but it is difficult to set up a reward system for negative results.

A lot of effort is being put into predicting the long-term stability of nuclear devices. Any of that technology which can be spun out would be extremely valuable for industry in general.

In chemistry departments, most modeling is currently quantum mechanics. Most modeling is done in chemical engineering departments. Engineers and chemists think about problems in different ways. There needs to be a better balance in funding between the fundamental science and the engineering applications. Some of the most inventive algorithms are being developed by engineers these days, but I would like to see more chemists get involved with that.

On Expectations of Commercial Software

Producing commercial quality code is ten times more difficult than writing academic code due to the necessity for error checking, robustness, and quality testing. This is not well understood by customers and academics.

It is natural then that cost is an issue. If we had a large market for the software, the price would plummet. We have a mismatch between the rate of adoption and the cost and effort to produce the codes. That's what keeps the price up. By going to PC-based platforms, the

hope is we'll get larger adoption, and because of that the price will come down and everyone will be a lot happier.

For similar reasons, it can be difficult to respond quickly to customer requests. If you allow rapid response to customer wants in particular areas, it is unfortunately like writing a report by committee that has to be collated at the end. The code will need to be rewritten and integrated at the end of the day. If we did that, the costs would skyrocket. The reality is that doing this cannot be managed effectively.

People do not have the same expectations of Microsoft Word. How are we different? Modeling is not a mature field like word processors. One word processing application is not that different from the next, but one catalysis problem is vastly different from the next.

On Employment Possibility and the Status of Modelers

Right now it's my impression that industrial positions in modeling are difficult to fill. There's a shortage of people. In the polymer field there are actually relatively few academics who are doing polymer modeling these days.

Another point—in physics—modeling and simulation are now considered equal to experiment in the APS. Is that true in the ACS?

On the Future of Molecular Modeling as a Bench Tool

The average synthetic chemist is quite capable of understanding the quantum tools. They [synthetic chemists] know enough about electronic structure to interpret the results. They need to be able to do more quantitative predictions with methods they can trust for it to be a boon. This hasn't happened yet because there remains an academic prejudice about understanding the tool you're using. To measure UV spectra, does one need to know how to build the spectrometer? No. There is still that prejudice about the calculations—that one shouldn't be doing quantum calculations unless you know how the basis set is constructed! This type of prejudice tends to evaporate with time, but it still exists. So today, one can get meaningful results on a UV spectrometer without having to build one. Perhaps it is the responsibility of the software company and the government to come up with computational standards. We need an ASTM for quantum mechanical methods! If that were done, we would find a lot more synthetic chemists making use of these tools. [Editor's note: See NIST site report for one proposed approach.]

If typical inorganic chemists were to use these tools, they would be delightfully enlightened. (Plus, there is the added benefit of having to devise a strategy rather than shake and bake.) One of the nice things about inorganic chemistry is that the number of reactions are really very limited—electron transfer, proton transfer, exchange, addition and elimination. A limited number of tools are required which could address nearly the complete set of problems.

REFERENCES

Brooks, B.R., R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus. 1983. *J. Comp. Chem.* 4, 187-217.

Richon, A.B. 1996. Mergers and alliances within computational chemistry. Network Science Corporation. (http://www.netsci.org/Science/Compchem/feature17.html).

Site: **Motorola, Inc.**
Semiconductor Products Sector
Computational Materials Group
3501 Bluestein Blvd.
Austin, TX 78721
http://www.motorola.com

Date Visited: 21 January 2000 in Albuquerque NM

WTEC Attendees: R. Miranda (report author), E.B. Stechel, P. Vashishta, P.R. Westmoreland

Hosts:
Murray Daw—At the time of the visit: Senior Scientist and Section Leader, Digital DNA Laboratories murray.daw@motorola.com, 512-933-6016, 512-933-6330. PhD Physics, four years at Sandia National Labs, 18 months at Motorola, based in Austin. Presently Professor of Physics at Clemson University.
Roland Stumpf, roland.stumpf@motorola.com, 505-845-7898. PhD Physics, surface science (Eindhoven), post-doc Washington U. and Sandia National Laboratories, 2 years at Motorola. Based at Sandia.
Wolfgang Windl, wolfgang.windl@motorola.com, 512-933-2307. PhD Theor. Solid State Physics, post-doc Arizona St. U. and Los Alamos, 2 years at Motorola. Based in Austin, but currently at Los Alamos.
Benjamin Liu, Motorola SPS-CMG-Los Alamos, benjamin.liu@motorola.com, 505-667-0860, fax 505-665-1999. PhD Physics, surface science (U. Penn.) and materials science (U. Illinois), post-doc Oak Ridge, 2 years at Motorola. Based at Los Alamos.

BACKGROUND

Motorola is a $30 billion, 150,000-employee worldwide organization currently headed by Chris Galvin III. Motorola's beginnings were in 1928 when Paul and Joseph Galvin incorporated the Galvin Manufacturing Corp. in Chicago, IL. In its 70-year history Motorola has pioneered many advances in the semiconductor, electronics, and information technology sector. Developer of leading automobile radios, color televisions, walkie-talkies, pagers, cellular phones, it is currently one of the leaders in wireless digital communication. In the mid-90s it started a cellular communications system called Iridium, designed to reach every point on the globe and based on an array of 66 small satellites in low-earth orbit. More recently, to populate the market of wireless messaging and multi-media products, Motorola has formed the Messaging, Information and Media Sector.

Semiconductor manufacturing (in the 1950s and '60s), and integrated circuit manufacturing since the '70s, have always been among Motorola's core strengths. The Semiconductor Products Division was established in the '70s under President Robert Galvin. The Center for Simulation and Modeling of Semiconductors originated in the early '90s as a CRADA with Los Alamos National Laboratory. Then in 1997 Motorola established the Computational Materials Group (CMG) within the Predictive Engineering Laboratory. Currently, the CMG under the DigitalDNA Laboratories has established another CRADA with Los Alamos Computational Materials Laboratory.

Motorola is organized into two primary design and manufacturing sectors: the Semiconductor Products Sector (SPS) and the Integrated Electronics Systems Sector. Its business units are organized under a single umbrella, the Communications Enterprise.

As the world's number-one producer of embedded processors, Motorola's Semiconductor Products Sector offers multiple DigitalDNA solutions for the consumer, networking and computing, transportation and wireless communications markets. In addition, it provides high-volume of discrete, analog and digital semiconductors. The SPS produces $7 billion in annual revenue and employs 45,000 persons. The DigitalDNA Laboratory, created in 1999, is part of SPS and is responsible for its research and development.

The Integrated Electronic Systems Sector designs and manufactures a broad range of electronic components, modules and integrated electronic systems and products for automotive, computer, industrial, transportation, navigation, energy systems, consumer and lighting markets.

The Communications Enterprise aligns all of Motorola's communications businesses into a coordinated unit of major business sectors: Personal Communications, Broadband Communications, Commercial, Government and Industrial Solutions, and Network Solutions.

MOLECULAR MODELING R&D ACTIVITIES

Computational materials research at Motorola is done at its corporate headquarters (Physical Sciences Research Lab, Illinois) and at its business units in Austin and Phoenix. Our hosts represented the views of the SPS-DigitalDNA Computational Materials Group that is based in Austin and Phoenix, with presences at Los Alamos and Sandia National Laboratories.

Their goal is to solve materials problems encountered during IC manufacturing and to improve process and device modeling for engineering applications. They bring in advanced solid-state theories and provide fundamental understanding.

The CMG has six physicists and materials scientists, four continuum modelers and two engineering modelers. Seven members are located at Los Alamos (five under Motorola payroll and two under LANL), four in Austin and one in Phoenix.

Materials modeling activities are divided into front-end and back-end applications. Front-end includes the Si, source and drain dopants, gate, and gate oxide. The back-end includes wiring and dielectrics. Both involve atomic-scale calculations, used to elucidate diffusion mechanisms.

In front-end applications, the challenge is to control tightly the dopant profiles in Si, which implies good understanding of implantation and diffusion mechanisms. Commercial TCAD software can be used. The CMG's task is to improve diffusion and clustering models and parameters, to produce new models for the implantation damage and the effects of stress on diffusion.

In the back-end applications, the challenge is the variety of materials utilized, e.g., Cu, magnetic layers, high-K gate oxides, low-K dielectrics like porous and fluorinated Si, and Si-Ge. CMG's activities are to study the effectiveness of diffusion barriers (e.g., for O diffusion in Pt [1]), the intermixing of layered materials (e.g., in magnetic RAM), and model deposition techniques (e.g., ionized PVD for Cu).

Another significant activity is the study of quantum transport, which has become important with tighter devices. The challenge is to produce models that are computationally much faster than the state of the art.

The CMG maintains a 220-CPU Linux cluster (110 nodes x 2 CPU/node) with 1 GB of RAM per node. The interconnect is performed with 5 switches x 100 Mbps Fast Ethernet backplane. Each job is allocated 30-40 CPUs, thus extensive parallelization is used. Currently, the CMG uses the following methods:

- Electronic DFT, with the Vienna package (VASP)
- Atomistic simulation of solids
- Semi-empirical MD; embedded-atom method used in special cases
- Non-equilibrium Green's Function Method, developed in-house for quantum transport

DISCUSSION

For Motorola, emerging applications for molecular and materials modeling are in manufacturing processes; mechanical properties; and electronic, optical, magnetic and chemical properties, including the following:

- Implantation and diffusion of dopants in Si
- Deposition of various materials in thin layers, including growth mechanisms and control
- Stress, permeability, deformation, and microstructure of thin films
- Electronic and transport in small devices
- Magnetic materials
- Adhesion and compatibility with organic compounds, such as photoresists, organic semiconductors, and biological molecules

Existing methods are helpful but need advancement. Local density approximation DFT has serious limitation because of the known band-gap and valence-band width errors. Improved determination of active diffusion paths is needed, as, for example, with the Directed Dimer method. Accelerated MD, such as hyper-MD, is essential for performance calculations. The broad applicability of materials modeling will remain limited by the shortcomings mentioned, in particular by the inability to extend MD to long time-scale.

Collaborative research groups between industry and national laboratories can help accelerate progress and extend applicability by developing and coding advanced physical models. For example, current commercial software for implantation and diffusion simulations rely on overly simplified models, which must be recalibrated at each generation to be effective.

Advancements in theory also help to reduce experimental costs, by providing adequate guidance, for example, in the selection of materials. The reverse is also true. Collaboration with national laboratories and universities to perform limited experimental testing proves helpful.

In the distributed modeling environment practiced by the CMG, proper identification of company priorities is challenging but essential for effective coupling. New projects are usually the result of specific questions from the business units, as well as long-term strategic research to maintain competitive advantage. Knowledge of the modeling capabilities in-house usually leads to new and significant questions, hence the importance of educating the company. Afterwards, the interaction of modelers with early experimental work on a project is highly beneficial. When the project is underway, weekly meetings and communication of results is found necessary.

Publication of results is encouraged at the CMG. In the particular case of this CRADA, the CMG enjoys the opportunity of being able to publish most of their research, which helps attract and maintain scientific talent. However, integration with older company styles is difficult and poses questions about continuity of research after the CRADA.

CONCLUSIONS

The CRADA between the CMG at Motorola DigitalDNA Laboratories and the Computational Technologies Laboratory at Los Alamos National Laboratory is an example of effective industry-government cooperation. A young team of talented modelers has made significant contributions to Motorola's core competencies. Simultaneously, they are increasing the scientific value of current molecular modeling methods, particularly in the area of dopant diffusion and physical vapor deposition.

REFERENCE

Stumpf, R., C.-L. Liu, and C. Tracy. 1999. Reduced oxygen diffusion through beryllium doped platinum electrodes. *Appl. Phys. Lett.*, 75(10) 1389-91.

Site: **National Institute of Standards and Technology (NIST)**
100 Bureau Drive
Gaithersburg, MD 20899
http://www.nist.gov/

Date: 20 February 2000

Report authors: K.K. Irikura and S.C. Glotzer

BACKGROUND

NIST, formerly the National Bureau of Standards (NBS), was established in 1901 as the first U.S. government research laboratory and remains a highly respected research institution. The change of name from NBS to NIST occurred in 1988, at the same time that its mission was amended to include aiding American industry in commercializing technologies. NIST is part of the U.S. Department of Commerce, and its Director is appointed by the President. It is best known as the source of primary measurement standards, including fundamental units (e.g., mass, length, physical constants), well-characterized materials for calibration purposes, and critically evaluated scientific and engineering data. NIST has a total of 3300 staff and 1500 visiting researchers at its campuses in Gaithersburg, Maryland and Boulder, Colorado. In addition to the research laboratories, NIST operates the Advanced Technology Program (ATP), the Manufacturing Extension Partnership (MEP), and the Baldridge National Quality Program. The total budget is about $800 million for FY 2000.

MODELING R&D ACTIVITIES

There are seven research laboratories within NIST: Electronics and Electrical Engineering, Manufacturing Engineering, Chemical Science and Technology, Physics, Materials Science and Engineering, Building and Fire Research, and Information Technology. Most of the activities in molecular or materials modeling are in three laboratories as summarized below. NIST central computing facilities include an 80-CPU IBM SP2 and four SGI Origin 2000 machines (104 CPUs).

Chemical Science and Technology Laboratory (CSTL)

Modeling efforts in CSTL are concentrated in the Center for Advanced Research in Biotechnology (CARB, http://www.carb.nist.gov/) and the Computational Chemistry Group (http://www.nist.gov/compchem/). CARB is a joint institute between NIST, the University of Maryland, and Montgomery County, Maryland. Modeling at CARB (12 staff, including three permanent faculty members) aims to understand the structure and function of nucleic acids and enzymes. Most of the associated computations are based upon force fields, but *ab initio* electronic structure calculations are also important. In contrast, research in the Computational Chemistry Group (five staff, including four permanent staff members) uses quantum chemistry methods almost exclusively. They focus on physical chemistry applications such as chemical kinetics, thermochemistry, and molecular spectroscopy. About five additional staff in the adjacent Experimental Kinetics and Thermodynamics Group spend significant time doing *ab initio* modeling. Research both at CARB and the Computational Chemistry Group includes a mixture of applications and methods development. Both CSTL groups maintain databases whose customers include researchers in industry, academia, and other government laboratories.

Materials Science and Engineering Laboratory (MSEL)

Materials theory, modeling and simulation has long played an important role within all five divisions of MSEL, including Polymers, Metallurgy, Ceramics, Materials Reliability, and the Reactor Division (now the NIST Center for Neutron Research or NCNR). Traditionally, most simulation in MSEL has focused on the mesoscale, using, e.g., Cahn-Hilliard/time-dependent Ginzburg Landau methods to model phase separation of alloys and polymer blends, and on the macroscopic scales, using finite element methods to model, e.g., flow of polymers during injection molding and fracture of ceramics. In the NCNR, researchers use

molecular dynamics methods to study protein dynamics and other problems in complex macromolecular fluids to help interpret neutron scattering experiments.

In recognition of the growing importance of modeling and simulation in materials science and engineering, in 1994 MSEL established the Center for Theoretical and Computational Materials Science (CTCMS). The mission of the CTCMS is to advance progress in computational materials science, and accelerate its use by, and impact on, industrial R&D. The CTCMS (http://www.ctcms.nist.gov) has a director (Sharon Glotzer) and a deputy director (James Warren), both computational condensed matter/stat-mech physicists by training, and a (nearly) full-time system administrator. The CTCMS works with the divisions in MSEL to identify new opportunities in materials theory and simulation important to U.S. industry, initiate new activities, assemble critical mass in the form of postdocs, guest researchers, and internal and external collaborators, and establish the state of the art in the field. The CTCMS typically holds workshops as the first step in kicking off a new activity. Teams of researchers, or "working groups," consisting of both computational materials scientists and experimentalists from industry, universities, NIST, and other government and national laboratories are formed to attack materials problems best addressed through the collaboration of multiple researchers with different backgrounds and expertise. The CTCMS provides small amounts of funding ("glue money") to support working group meetings, travel, and a postdoc or graduate student or two. An MSEL staff member (usually but not always a computational materials scientist) typically serves as one of the PIs and coordinator of the activity. The CTCMS operates a highly modern central facility within the Materials building that provides computational resources and physical workspace for CTCMS members. The operating budget for the CTCMS is $0.5 million in FY2000 (not including staff salaries, major equipment, or cost sharing), reflecting an increase over previous years.

Current CTCMS working group activities in molecular and materials modeling and simulation include deformation of metals; development and validation of micromagnetic modeling tools; protein crystallization; Green's functions and boundary element modeling tools for materials; atomistic and multiscale molecular dynamics simulations of filled and nano-filled polymers; glass formation and transport in cold liquids; first principles prediction of phase diagrams for ceramics; development of modeling tools for predicting microstructure of blends under shear; and the development of object oriented modeling tools for predicting materials properties from microstructures. The CTCMS supports the development of public domain modeling and simulation research tools, and disseminates these tools on its Web site. In the past 12 months alone, CTCMS-developed modeling tools were downloaded nearly 3000 times by researchers around the world. One of these tools, known as OOF, recently won the *Industry Week* award for being one of the top 25 new technologies of 1999. While most CTCMS activities directly involve industrial researchers, only one CTCMS activity has involved a CRADA (cooperative research and development agreement). The CTCMS also solicits proposals for workshops from non-NIST researchers, and has hosted and/or sponsored close to 30 workshops on computational materials science since 1995. Many of these workshops have launched new collaborative activities in computational materials science, and all have brought researchers from industry together with researchers from academia and government and national labs.

Building and Fire Research Laboratory (BFRL)

The Fire Sciences Division within BFRL has a small program (one permanent staff member, Marc Nyden) employing both quantum and molecular mechanics to address issues in polymer combustion and charring. This is done in close collaboration with experimental research in the division. In additional, one permanent staff member spends a portion of his time developing and using lattice Boltzmann and dissipative particle dynamics methods for modeling fluid flow of multiphase and/or multicomponent fluids, and suspensions. Most of this work is done in collaboration with researchers in other laboratories.

Many of the modeling and simulation activities within the three laboratories above involve collaborations with computational scientists in NIST's Information Technology Laboratory (ITL). Contributions by ITL include helping with parallelization of modeling codes, development of graphical user interfaces for research simulation tools, and development of simulation code.

DISCUSSION

The rapidly decreasing cost of computing, combined with the generally rising costs of experimental research and testing, has led to increasing use of modeling throughout the R&D community. This is reflected in the NIST laboratories.

The NIST research environment encourages more collaborations than are typical within a university. There are many examples involving molecular modeling and simulations. *Ab initio* results from CSTL are used in BFRL and elsewhere in CSTL to build high-level computational models of combustion processes. These in turn are connected to experimental investigations of flame inhibition and of atmospheric chemistry. Other CSTL modeling is done in collaboration with experimental spectroscopy in the Physics Laboratory. Computational and experimental scientists work together at CARB on problems such as protein folding. Modeling and simulation are well integrated with experimental programs in MSEL for a number of wide-ranging problems, including solder interconnect design, magnetic materials, fracture of ceramics, filled polymers, phase separation of blends, solidification of alloys, and prediction of phase diagrams for ceramics.

Molecular and materials simulation at NIST is supported at the highest levels, and it is continuing to receive substantial recognition and support from NIST management.

Site: **Owens Corning, Inc.**
Owens Corning Science & Technology Center
2790 Columbus Road; Granville, Ohio 43023-1200
http://www.owenscorning.com/

Date of interview: 13 December 1999

Reporter: P.R. Westmoreland

Contact: Norman T. Huff, tom.huff@owenscorning.com

BACKGROUND

Owens Corning is a global company that provides insulation, roofing, other building materials, and composites to companies and individual consumers. It has been a leader in the making of glass fibers. In 1999 its sales were $5.05 billion and profits were $270 million. However, on October 5, 2000, it filed for Chapter 11 bankruptcy protection due to lawsuits concerning asbestos-containing insulation products that it had discontinued in 1972.

MODELING R&D ACTIVITIES

Owens Corning began work in molecularly-based modeling in 1995, focusing on needs to understand silicate glass structures. Subsequently, the work began to focus on interaction of organic molecules with silicate glass surfaces.

Dr. Huff worked in this area full time for three years, and three others worked a small amount of time. At present there is no one working in this area. Cerius2 from MSI and Gaussian were the main programs. Computers included an SGI Origin 2000 (6 processors) and SGI dual processor Octanes.

Presentations have been made regarding the factors affecting the generation of glass structures. Force fields for silicate glasses were developed in conjunction with William Goddard at Caltech. They showed that prior molecular dynamics work in generating glass structures has not adequately investigated the effect of cell size and cooling method upon the final structure obtained.

To study the interaction of molecules with glass surfaces adequately, Owens Corning researchers concluded that quantum MD is required. Because of the random structure of silicate glasses, very large numbers of heavy atoms must be considered in any calculations. A second best approach would be reactive force fields that would be suitable for MD simulations.

In general, Dr. Huff and his colleagues see the biggest needs as more capable computers and quantum algorithms that scale close to N.

Site: **Pharmacopeia, Inc.**
CN 5350
Princeton, NJ 08543-5350
(609) 452-3600; Fax: (609) 452-3671
http://www.pharmacopeia.com/

Date Visited: 17 January 2000

WTEC Attendees: P.R. Westmoreland (report author), P.A. Kollman, K. Morokuma, E.B. Stechel, P. Vashishta

Hosts: John J. Baldwin, Chief Science and Technology Officer (baldwin@pharmacop.com)
Peter Gund, Senior Director, Molecular Modeling (MM) (pgund@pharmacop.com)
Kenneth M. Merz Jr., Senior Director, Center for Informatics and Drug Discovery (CIDD) (kmerz@pharmacop.com)
Janet Cohen, Software Development Group (jcohen@pharmacop.com)
Tariq Andrea (MM), Ailan Cheng (CIDD), David Diller (CIDD), Steve Dixon (CIDD), William J. Egan (CIDD), Evert J. Homan (MM), Giorgio Lauri (CIDD)

BACKGROUND

Pharmacopeia Inc. was formed in 1993. In June 1998, it acquired Molecular Simulations Inc. from Corning. The combined company, structured as the parent of Pharmacopeia Laboratories (Princeton NJ) and of MSI, reached revenues of $100 million in 1999. It specializes in modeling software and in drug discovery by its own tools of combinatorial chemistry, high-throughput screening, molecular modeling, and informatics.

The first subsidiary, Pharmacopeia Laboratories, is a 220-employee company based in Princeton NJ which focuses on accelerating the drug discovery process. Pharmacopeia Labs has collected or synthesized 6.2 million tagged chemical compounds for composing problem-specific combinatorial-chemistry libraries, adds about 1 million compounds per year, has developed a 1536-well assay plate and supporting equipment for high-throughput screening (compared to standard 96-well plates), develops new bioinformatics methods, and uses these tools in various types of collaborations and contracted development.

MSI is a leading software house in the area of molecularly based modeling for chemistry, materials, and life sciences. MSI's 330 employees are located primarily in Cambridge UK and in San Diego CA. (The site report for MSI San Diego, presented separately, emphasizes more of the non-life-sciences side of MSI/Pharmacopeia software.)

[Note: Later in 2000, Pharmacopeia added additional software strength by its purchase of the software subsidiaries of Oxford Molecular Group (bioinformatics and cheminformatics) and Synopsys Scientific Systems Ltd. (bioinformatics). The software subsidiaries were combined under a new name, Accelrys as of June 1, 2001.]

R&D ACTIVITIES

Dr. Baldwin presented an overview of the entire company; the subsequent presentations focused on activities at Pharmacopeia Laboratories. With technology and services as one aspect, and discovering potential new drugs as another, Pharmacopeia seeks to build on its platform technologies of chemistry, biology, and software to work in four business areas:

1. Software sales and service
2. Chemical compound leasing
3. Drug discovery services
4. In-house drug discovery

MSI is at the heart of the first business. Pharmacopeia representatives stated that MSI software is in most or all of the major pharmaceutical, chemical, petroleum and gas companies. All of MSI's work has recently been regrouped as being "life sciences" or "materials science," the second of which includes chemicals and fuels along with hard and soft condensed matter.

Pharmacopeia Labs handles activities in the other three business areas. Chemicals are synthesized and linked to tagged solid beads for rapid identification, 500 picomoles per bead, and the beads are physically split 50-fold. The large inventory of chemical compounds is then selected into custom libraries leased for testing by companies currently including Bayer, Novartis, Schering-Plough, and Zeneca. Alternatively, Pharmacopeia Labs can do the testing under contract, either to the extent of initially identifying or optimizing drug candidates for interacting with specified proteins. Self-funded drug discovery is also going on. Success has been demonstated in each of these three business areas through subsequent identifications of activity in animal models.

Computer modeling in Pharmacopeia Labs includes molecular modeling (led by Dr. Gund), "pharmaconomics" (in the Center for Informatics and Drug Discovery, led by Dr. Merz), and computer science support (led by Janet Cohen).

Molecular modeling here is largely to assist the synthetic chemists and biologists. Originally, it was used most by the bench chemists, but it increasingly was left to the specialists. Now, due to the MSI Weblab interface, the results and even appropriate calculations have become more accessible to the chemists. Most modeling is molecular simulations, with computational quantum chemistry mainly on the side. Nevetheless, there was strong interest in new, more accurate force fields based on quantum calculations.

Professor Kollman noted that Dr. Gund had been at Merck before he came to Pharmacopeia, where he was a pioneer in successful molecular modeling for pharmaceutical purposes. Dr. Gund noted that at Merck and many other companies, the future of drug development had once seemed to be rational design of drugs by computation, then combinatorial chemistry, but neither was enough by itself.

CIDD is a relatively new group, created in November 1998 to deal with the explosion of data from the genomics and high-throughput screening areas. Its activities are divided into four parts:

1. Library design (Lauri)
2. Databases (I-Ping Cheng)
3. Docking (Diller and Dixon)
4. Data mining and bioavailability (Egan and A. Cheng)

Pharmacopeia researchers note that success depends on encompassing desired activity, solubility, bioavailability, and toxicity. At this point, they have already scored successes for a set of kinase targets.

Computer science support is both at Princeton and with the MSI half of the company. They have developed a Pharmacopeia Design Environment to develop virtual libraries and then to narrow the list of compounds, moving to consider issues like bioavailability. For data mining, they have developed a Pharmacopeia Information Environment using Oracle and Daylight-based tools, the databases, and Visual Basic. Database information is typically stored as SMILES strings.

Most of the molecular modeling work is carried out on SGI R10000-class workstations, and CIDD heavily uses PCs and a Sun server for information databases. They expect to see increased use and usefulness of parallel computing. MSI and internally developed software is used for much of the work, but external software is also used to a significant extent.

DISCUSSION

Prior to the purchase of MSI, Pharmacopeia's molecular modeling had been mainly with Tripos software. Other companies were considered, but the company and its customers recognized considerable potential for

cross-links and synergy between Pharmacopeia's and MSI's strengths. Synergies between modeling, experiments, and informatics have been especially useful.

Close work with the chemists, including joint meetings, is very important, proving to be effective at technical advances and at aiding long-term vision (versus the "what have you done for me today?" problem). Giorgio Lauri emphasized that another effect is the value of understanding enough about a problem before plunging deeply into calculations. Bringing hands-on tools into the experimentalist-modeler relationship is convincingly effective.

When a company is so physically spread out (Princeton, Cambridge, San Diego), collaborations can be aided by electronic conferencing. The company maintains video conferencing facilities connected by a dedicated line.

This group is confronting many of today's technical needs, but other needs will become important. The Human Genome Project will make functional genomics increasingly important, leading to more predictive target and drug structures. ADME modeling will become more sophisticated and useful, as will formulation modeling. Principles need to be better understood to give better docking and scoring. They recognize the limitations of current force fields and would like to see more fundamentally sound force fields developed.

Other needs are less directly technical. Better preparation of scientists would include more practical statistical training, attention to oral as well as written communications skills, and development of rigorous record-keeping skills. The last is a complicated one for modeling because it is impractical to include printed or screen output in written, bound lab notebooks. However, it really is not so different from the inability to include complete data recorder output or to record every visual observation photographically. In both cases, one should write why and what you're doing, what you find, and how you interpret it. For students, this group would recommend more internships and pre-doctoral jobs with companies. Such preparation also helps students learn what it really means to work across disciplinary boundaries.

CONCLUSIONS

A vigorous young company is applying broad visions of what can be done in drug discovery with the aid of computers. The MSI portion of the company complements the drug discovery effort. However, MSI has such a large presence in the non-biological chemicals and materials sectors that molecular modeling in these sectors could be neglected and harmed (a fear expressed at several sites we visited) if MSI is driven too much by life sciences. The challenge for Pharmacopeia is to balance these visions. At the same time, the chemical processing industries project an increasing role for life sciences and bioanalogs. The synergism at Pharmacopeia could then have very positive impacts for this sector.

Site: **Phillips Petroleum Company/Chevron Phillips Chemical Company**
Phillips Research Center
Bartlesville OK 74004
http://www.phillips66.com/

Date: 10 February 2000

Reporter: P. R. Westmoreland

Contact: Donald E. Lauffer, 205 CPL PRC, 918-661-9724, Fax 918-662-1097, del@ppco.com
[Note that the report is based on public sources and on Phillips-approved material received for this report from Donald E. Lauffer and Donald G. Truhlar, not on a site visit or on interviews.]

BACKGROUND

Phillips Petroleum Company is an oil, gas, and chemicals company that carries out integrated activity in resource exploration, production, refining, chemical manufacturing, transportation, and marketing. Worldwide, it employs 12,143 people plus 3,400 legacy Phillips CPC employees (as of 8 Aug 2000). Cash flow from operations was $1.9 billion in 1999, yielding $548 million net operating income of which $146 million was from chemical operations.

On July 1, 2000, Chevron Phillips Petroleum Company (CPC) was created as a 50/50 joint venture with Chevron Corporation, bringing together 3,400 people from the chemicals business of Phillips and 2600 people from Chevron's petrochemicals and plastics businesses. The company's products are olefins, aromatics, polyolefins and other polymers, and specialty chemicals.

CPC has 400 scientists and other researchers in five locations: Richmond, CA, and Kingswood, TX (petrochemicals); Bartlesville, CA (polymer research and development); Marietta, OH (styrenic polymers); and Orange, TX (polyethylene and technology).

MOLECULAR MODELING R&D ACTIVITIES

In recent years, Phillips has had great success with a new linear low-density polyethylene (mPACT®, introduced in 1998) made with a proprietary metallocene catalyst. LLDPE production capacity located at Cedar Bayou, Texas, reported in the company's 2000 Annual Report, was 400 million lb./yr.

Reported work on design of polyolefin catalysts seems clearly related to such activities. In 1995, the NIST ATP program co-funded a project with Phillips on "Computational Methods for Catalyst Design" (http://jazz.nist.gov/atpcf/prjbriefs/prjbrief.cfm?ProjectNumber=95-05-0012). The stated objective was to "develop and apply computational chemistry methods that will accelerate the discovery of metallocene-type catalysts for ethylene polymerization." Project funding was $3.474 million with $1.705 million from ATP. Participants included the University of Minnesota Supercomputing Institute, Emory University, and the University of Oklahoma.

In the project, computational chemistry methods were developed by researchers at Phillips and its university partners to accelerate the discovery and evaluation of new single-site catalyst systems for ethylene polymerization. The computational methodology was based on a fundamental understanding of the molecular structure of metallocene polymerization catalysts and a rationale for the key steps in the polymer formation mechanism. The project utilized both computational chemistry and experimental catalyst kinetics to develop methods to predict polymerization reactivity.

Electronic structure calculations (both gas-phase and liquid-phase) utilizing density functional methodologies were employed to predict the energetics and geometries of reactant, intermediate, transition, and product states. The selected methodologies were applied to structurally diverse metallocene catalyst systems. The

metallocenes were zirconium-based with cyclopentadienyl, indenyl, and fluorenyl ligands, either non-bridged or with C_1, C_2, and Si-bridges and with methyl and n-butyl substituents.

More details were presented in two papers by Das et al. (1998 and 1999) entitled "Ethylene Polymerization by Zirconocene Catalysis." These papers describe the computational collaboration involving Phillips, the University of Minnesota Supercomputing Institute, and Emory University.

Bound species and transition states were modeled for interactions of ethylene, propylene, and 1-hexene with each of two zirconocenium systems, Cp_2ZrR^+ and $[CpCH_2Cp]ZrR^+$. The B3LYP density-functional method was applied using the Los Alamos LANL2DZ effective-core-potential (ECP) basis set for atoms directly involved with the reaction and the next nearest atoms, while a 3-21G basis set was used for other atoms. LANL2DZ with a relativistic ECP basis set was used for Zr. The semi-empirical molecular-orbital code AMPAC (Semichem, Inc.) and the density-functional code DGAUSS (Oxford Molecular Group, UK) were adapted to include the SM5.2R solvation model of Cramer and Truhlar. Gaussian 94 (Gaussian, Inc.) was also used, in particular with Morokuma and co-workers' ONIOM method for extrapolation of *ab initio* methods to large molecules.

Differences in gas-phase and solvated metallocene were significant, and the differences for particular steps pointed toward important selectivity differences. The olefin insertion step was identified as the probable rate-determining step, and the most probable chain termination step was identified.

A key result is that some of the polymerization steps had differing calculated activation energetics that were dependent on the molecular structures of the metallocenes. Relative performance properties of the metallocenes were predicted from these energetics.

In further work, a series of structurally diverse metallocenes were synthesized and purified. The influences of specific reaction parameters on polymerization rates and polymer properties were experimentally measured for ethylene polymerizations by the metallocene catalyst systems. Correlations of these experimental data with the computational results for these catalysts indicate that systematic relationships exist between the structure-related properties and quantities involving the calculated activation energies of metallocene catalysts. Catalyst performance properties that can be determined include catalytic polymerization activity, the propensity to incorporate alpha-olefin co-monomers, molecular weights of produced polymer, the influence of hydrogen on polymer molecular weight, and ability to self-induce short-chain branching.

The computational methodology enables chemists at Phillips to generally predict key catalyst performance properties based upon the molecular structure of a metallocene. Effective use of computational chemistry at the conceptual stage of catalyst discovery facilitates a better understanding of the fundamental catalyst chemistry, and its use for screening and prediction reduces the number of experimental catalyst evaluation studies by computationally screening catalysts for desired properties.

Software developments from this project have been incorporated into programs made available to other computational chemists. In particular, various resulting code developments were made available, or are in the process of being made available in AMPAC version 5.4ml (distributed by Semichem, Inc.), DGSOL version 4.0ml (distributed by Oxford Molecular Group, UK), and GAUSSIAN (distributed by Gaussian, Inc.). Further details on the new versions of AMPAC and DGSOL are available at http://comp.chem.umn.edu/truhlar/.

CONCLUSIONS

The Phillips/NIST ATP project was focused on developing useful models and methodologies for such practical problems as homogeneous catalyst design. The Executive Summary of the final report from that project (Lauffer et al. 1998) stressed the value to the industrial partner of using these approaches for screening and prediction at the early stages of development.

Modeling of homogeneous catalysts has proven to be a crucial feature in developing new metallocene catalysts and polymers for several companies worldwide, for example Union Carbide, Mitsubishi, and EniChem. This ATP project is thus an excellent example of how government can aid important developments in modeling that benefit the whole community, while at the same time providing an incentive for a co-funding company through hands-on experience that can be applied to future problems.

REFERENCES

Das, P.K., D.W. Dockter, D.R. Fahey, D.E. Lauffer, G.D. Hawkins, J. Li, T. Zhu, C. J. Cramer, D.G. Truhlar, S. Dapprich, R.D.J. Froese, M.C. Holthausen, Z. Liu, K. Mogi, S. Vyboishchikov, D.G. Musaev, and K. Morokuma. 1998. Ethylene polymerization by zirconocene catalysis. In *Applying Molecular and Computational Chemistry*, New York: American Institute of Chemical Engineers; presented at Annual Meeting of AIChE, Miami FL, November 15-19, 1998.

Das, P.K., D.W. Dockter, D.R. Fahey, D.E. Lauffer, G.D. Hawkins, J. Li, T. Zhu, C.J. Cramer, D.G. Truhlar, S. Dapprich, R.D.J. Froese, M.C. Holthausen, Z. Liu, K. Mogi, S. Vyboishchikov, D.G. Musaev, and K. Morokuma. 1999. Ethylene polymerization by zirconocene catalysis. In *Transition State Modeling for Catalysis*, edited by D.G. Truhlar and K. Morokuma (American Chemical Society Symposium Series Volume 721, Washington, DC), pp. 208-224.

Lauffer, Donald E., and Darryl R. Fahey (Phillips Petroleum), Keiji Morokuma (Emory Univ.), Donald G. Truhlar (Univ. Minnesota), and Ronald L. Halterman (Univ. Oklahoma), project managers. 1998. *Computational Methods for Catalyst Design*, Executive Summary. Advanced Technology Program of National Institute of Standards and Technology, Project 95-05-0012, 1995-1998.

Site: **Rohm and Haas Company**
Spring House Technical Center
727 Norristown Road
P.O. Box 904
Spring House, PA 19477-0904
Telephone: + 1-215-641-7000; Fax: + 1-215-641-7857
http://www.rohmhaas.com/

Date Visited: 19 January 2000

WTEC Attendees: P.T. Cummings (report author), P.A. Kollman, R. Miranda, K. Morokuma, P. Vashishta, P.R. Westmoreland

Hosts: Rafael G. Aviles, Research Section Manager
Susan Fitzwater, Senior Research Fellow, rs0sjf@rohmhaas.com
Mark R. Schure, Research Fellow
Chuck Reynolds, Research Fellow
Zhiwei Liu, Senior Scientist

BACKGROUND

Rohm and Haas Company is the third largest chemical company in the U.S., and the second largest specialty chemical company in the world. It employs over 18,000 people at 150 manufacturing and technical centers around the world. The company, with annual revenue approaching $7 billion, is the world's largest producer of acrylate and methacrylate monomers, and a major supplier of numerous specialty products for industry and agriculture. It recently acquired Morton International, producer of Morton salt and a wide range of specialty chemicals, which has diversified its product line. Morton International was roughly half the size of the Rohm and Haas company prior to the acquisition. Rohm and Haas groups its products into five main businesses: performance polymers, surface finishes, electronic materials, chemical specialties and salt.

The company invests about 5% percent of revenues in research and technology. Approximately 2,500 employees work in its "technology community" around the world, of which 25 percent hold PhD degrees. Most of the research is conducted near Philadelphia, PA. The 185-acre Spring House Technical Center is home to about 1000 scientists and support staff performing product-related research and development in support of the business units. The other principal research facility is Bristol Research Park, about 24 miles east of Spring House, hosting approximately 350 scientists and support staff. Additional research activities are conducted in Charlotte, North Carolina; Houston, Texas; Valbonne, France; Washinomiya, Japan; Singapore; Woodstock, Illinois; and at plant locations in Bristol, Pennsylvania and Lauterbourg, France. Agricultural field research is centered at experimental farms in Newtown, Pennsylvania; Waller, Texas; and Campinas, Brazil.

MOLECULAR MODELING R&D ACTIVITIES

The WTEC panel met with molecular modeling staff who focus on specialty chemicals, including agricultural chemicals. Additional activities may be pursued in the electronic materials area, but there is not significant linkage with the group visited by the WTEC panel. The molecular modeling group consists of five PhD chemists, one programmer and a UNIX support person. Tenure with Rohm and Haas is mostly long-term, and most of the members of the molecular modeling group had been with Rohm and Haas for most of their professional careers. The first molecular modeling staff member was hired in 1979 as a statistician. Currently molecular modeling is in a section with statistics and laboratory automation, and part of the Analytical/Computational Technology Center.

The markets served by molecular modeling include paintings and coatings, electronics, household products (detergents and personal hygiene), water treatment, agricultural chemicals, adhesives, and plastics. All the markets are very cost-sensitive, so that technical development has a strong cost focus.

There are three major kinds of work in the technical community at Rohm and Haas: synthesis (e.g., emulsion polymerization of acrylic monomers), process development (with the goal of getting consistency from every plant world-wide), and applications testing (e.g., performing the same tests on products as customers, such as peel test for adhesives).

Approximately half of the support for the molecular modeling group comes from selling services to technical customers via the business groups. The funding is usually for specific work on specific projects with short timelines and costs watched carefully. Some of the work of the group is funded by a shared services allocation to the businesses. This funding is usually used for evaluating and/or developing methods that should be useful in the future.

The range of molecular modeling projects includes QSAR (quantitative structure-property relationship), molecular and polymer modeling, chemical reaction modeling, and exploitation of new computational technologies (such as Web-based chemoinformatics). Because of the application in the paint and coating industry, there is a heavy emphasis on light scattering techniques, both experimentally and theoretically.

Molecular modeling has been used in patent activity at Rohm and Haas. In addition to being used to broaden patent coverages, modeling of chemistry has been used to defend a patent. Rohm and Haas proved in court that another company had infringed one of its patents on a urethane thickener. The essence of the Rohm and Haas position was that the process revealed in the patent of the other company could not avoid making the urethane thickener patented by Rohm and Haas, based on modeling. The code used to show this is now part of the court records. In other molecular-modeling-related patent activity, Thomas Pierce of Rohm and Haas has patented a computational methodology for modeling polymers (Blanco and Pierce 1993).

Examples of molecular modeling at Rohm and Haas include the following:

- Development of a model for paint formulation. Latex paint consists of polymeric binder, TiO_2 pigment and extenders (such as clay). There are thousands of different paint formulations. In order to develop a model for paint formulation, light scattering from TiO_2 needed to be understood and modeled. With this in hand, the company developed a molecular-based model for paint formulation that reduced the number of experiments required from thousands to about 300.
- Molecular modeling in support of the design of new insecticides. Rohm and Haas was awarded a 1998 Presidential Green Chemistry Challenge Designing Safer Chemicals Award for the invention and commercialization of a new chemical family of insecticides exemplified by CONFIRM™ selective caterpillar control agent (EPA 1998). The product, based on diacylhydrazines, acts by strongly mimicking 20-hydroxy ecdysone, a natural substance found within the insect's body which is the natural "trigger" that induces molting and regulates development in insects. The insecticide disrupts the molting process in target insects, causing them to stop feeding shortly after exposure and to die soon thereafter.
- Modeling of chromatography. The Theoretical Separation Science Laboratory at Rohm and Haas has made many contributions to understanding chromatography at the molecular level. A review (Schure 1998) and recent papers (Murphy, Schure, and Foley 1998; Martin, Siepmann and Schure 1999) exemplify the efforts. The latter work, involving the prediction of Kovats indices by Gibbs ensemble Monte Carlo simulation, is the result of a National Science Foundation GOALI award to the Theoretical Separations Science Laboratory of the Rohm and Haas Company and the University of Minnesota.

In the future, Rohm and Haas researchers foresee the following areas as needing improved molecular modeling tools: heterogeneous catalysis, mesoscale modeling of materials, physical properties of biomolecules, quantum chemistry for biomolecular systems, solvation, and colloids.

CONCLUSIONS

The molecular modeling activities at Rohm and Haas are well established and growing. The molecular modeling group is viewed as providing very high value services to the technical community. The group characterizes its relationships with its customers as "very good." By establishing its credibility over a long period of time, customers have become more and more willing to pay for molecular modeling activities.

REFERENCES

Blanco, M., and T.H. Pierce. 1993. US Patent Number 5260882: Process for the estimation of physical and chemical properties of a proposed polymeric or copolymeric substance or material.

Environmental Protection Agency (EPA). 1998. Press release about 1998 Presidential Green Chemistry Award (http://www.epa.gov/opptintr/greenchemistry/dsca98.htm).

Martin, M.G., J.I. Siepmann, and M.R. Schure. 1999. Simulating retention in gas-liquid chromatography, *J. Phys. Chem. B,* 103 11191-11195.

Murphy, R.E., M.R. Schure, and J.P. Foley. 1998. Quantitative analysis of poly(methyl methacrylate-butyl acrylate) copolymer composition by liquid chromatography-particle beam spectrometry. *J. Chromatography A*, 824, 181-194.

Schure, M.R. 1998. Particle simulation methods in simulation science. In "A Tribute to J. Calvin Giddings," edited by P. R. Brown and E. Grushka, *Advances in Chromatography* 39, 139-200, Marcel Dekker (New York).

Site: **Sandia National Laboratories**
PO Box 5800
Albuquerque, NM 87185
http://www.sandia.gov/

Date Visited: 21 January 2000

WTEC Attendees: E.B. Stechel (report author), R. Miranda, P. Vashishta, P.R. Westmoreland

Hosts:
Grant Heffelfinger, MS 0316, Manager, Computational Biology and Materials Technology Dept.; (505) 845-7801, FAX (505) 284-5451, gsheffe@sandia.gov
William J. Camp, MS Director, Computation, Computers, and Math Center, wjcamp@sandia.gov
Gary Grest, MS Materials & Process Computation & Modeling Department, gsgrest@sandia.gov
John Hamilton, MS 9161 Thin Film and Interface Sciences Department (by videoconference from Livermore, California), jchamil@sandia.gov
Elizabeth A. Holm, MS Materials & Process Computation & Modeling Department, eaholm@sandia.gov
James Tomkins, MS 1109, Computer Sciences Department, jltomki@sandia.gov

BACKGROUND

Sandia National Laboratories (SNL) began in 1945 on Sandia Base in Albuquerque, New Mexico, as Z Division, part of what is currently Los Alamos National Laboratory (LANL). Both LANL and SNL were established as part of the program to develop the atomic bomb in World War II—better known as the Manhattan Project. Sandia came into being as an ordnance design, testing, and assembly facility, and was located on Sandia Base to be close to an airfield and to be able to work closely with the military. In 1949, President Harry Truman wrote a letter to American Telephone and Telegraph Company President Leroy Wilson, offering the company "an opportunity to render an exceptional service in the national interest" by managing Sandia. AT&T accepted, began managing the Labs on Nov. 1, 1949, and continued in that role for nearly 44 years. On Oct. 1, 1993, the Department of Energy awarded the Sandia management contract to the Martin Marietta Corp., now Lockheed-Martin. Today, Sandia has two primary facilities, a large laboratory and headquarters in Albuquerque (more than 6,700 employees) and a smaller laboratory in Livermore, California (about 900 employees). Sandia is still a government-owned/contractor-operated (GOCO) facility.

SNL's original mission—providing engineering design for all non-nuclear components of the nation's nuclear weapons—continues today, but Sandia now also performs a wide variety of national security R&D work. The modified mission of Sandia is to assure a nuclear deterrent, to reduce U.S. vulnerabilities, to contribute to energy security and to respond to emerging threats.

Nearly 4,000 engineers (60%), scientists (24%), and other technical support staff (16%) make up Sandia's technical staff. By discipline the staff makeup is: 21% electrical engineering, 16% mechanical engineering, 9% physics, 8% computing, 5% chemistry, 3% mathematics, 23% other engineering, 7% other sciences, and 8% unclassified. By degree the staff makeup is: 41% masters, 35% PhD, 12% bachelors, 11% associate and 1% other. The annual budget for Sandia is $1.3 billion, with 72% allocated to national security, 9% to energy, 7% to environment, and 11% to other science and technology activities.

Sandia works closely with industry to design and produce components for nuclear weapons and is involved in several industrial consortia and industrial partnerships to develop and transfer technology. Sandia also provides support for more than 500 research projects at more than 100 universities in 32 states.

Sandia does not have a centralized modeling group. It does have molecular-scale materials simulation distributed across the two sites, in three divisions, at least four centers and a number of departments. There are several molecular-scale modelers in the following departments:

- Computational Biology and Materials Technology Department, which is in the Computation, Computers, and Math Center (New Mexico)
- Materials & Process Computation & Modeling Department, which is in the Materials and Process Sciences Center (New Mexico)
- Thin Film and Interface Sciences Department, which is in the Materials and Engineering Sciences Center (California)
- Several departments in the Physical & Chemical Sciences Center (New Mexico)

COMPUTATIONAL R&D ACTIVITIES

ASCI Red Computer System and CPlant Computer System

ASCI Red (http://www.sandia.gov/ASCI/Red/) is an Intel-based massively parallel supercomputer with 4,640 nodes configured with 1168 nodes dedicated for classified computing, 1168 nodes dedicated to non-classified computing and 2304 compute nodes that can be connected either to the classified or to the non-classified end. The computer was fielded in March 1996. Classified computing began in January 1997. The benchmark of 1.06 teraflops (TF) was demonstrated at Intel in December 1996. The full system was operational in June 1997 with a peak of 1.8+ TF and 1.338 TF on a Linpack Benchmark. After a 1999 upgrade, the peak increased to 3.15 TF with a 2.3796 TF Linpack Benchmark. The system architecture is distributed memory MIMD. The node architecture is two processors/node shared memory. The original processor was the 200 MHz Pentium Pro; the upgrade is the 333 MHz P II Xeon. The processor cache after the upgrade (2X across the board) is L1: 16 kB/16 kB and L2: 512 kB. The system RAM memory is ~1.2 TB.

Analysis of market conditions yields a number of implications for high-performance computing (HPC). The HPC market is small and shrinking (at least relative to other computing markets.) The performance of high-volume systems has increased dramatically. The Web is driving the market for scalable clusters. There is a hot market for high-performance interconnects. The implication of these factors is that high-volume building blocks are readily available. The commodity trends substantially reduce costs. Assembling large clusters is easier than ever. Incremental growth is possible. The market opportunity and solution that Sandia researchers are developing is commodity supercomputing: Cplant (http://www.cs.sandia.gov/cplant/). The basic concept is combining the following:

- Incremental growth
- Distributed between sites
- Specialized functionality
- Single-system capability

The CPlant technology and developed software has the "look and feel" of MPP; currently the only really scalable architectures are ASCI-Red and the T3E. The CPlant concept arose from "mentally disassembling ASCI-Red." As a consequence, CPlant is scalable to 10,000 nodes. Sandia researchers have demonstrated 125.2 GFlops with 350 EV56 nodes, which would place the machine at 53rd on the June 1999 top 500. With ~600 EV6 Nodes, they demonstrated 232 GFlops, which placed the machine at ~40 on the Nov. 1999 top 500, all for ~$10 million. That is definitely "bang for the buck." In a molecular dynamics benchmark consisting of 32,000 atoms/processor and testing up to 256 processors, the machine achieves nearly perfect scale-up and about half the performance of the ASCI Red Machine at considerably less cost. By mid-summer 2000, they expect to have a CPlant with a 1.5 TF peak (placing it in the top 10) for <$20 million. In contrast, ASCI-Red is $63 million (3.1 TF peak) and the two ASCI-Blue options are >$100 million each. They plan to achieve 100 TF for $150 million and petaflop performance before 2010.

History of Modeling and High-Performance Computing

Bill Camp credits Ken Wilson for origins of high-performance computing. In the mid 1980s, Wilson began talking about the third leg of science: computation & simulation on equal footing with theory and experiment. In 1987 Sandia fielded the first successful MPP machine: the 1024 nCUBE-10 MIMD

(independent processors) architecture. LANL fielded the CM2, which was bigger and based on a SIMD (single instruction) architecture. MIMD and SIMD competed for a while; MIMD survived. In the early 1990s, Sandia fielded the N^3-2 (1024 processors) with a peak of 2GF. Since 1993 Sandia has maintained an unbroken streak of fielding the largest supercomputer including the first terascale computer (ASCI Red) in 1996. In 1990, they successfully competed in the First National Grand Challenge Initiative—Materials by Design. In 1992, they identified the mesoscale as the stumbling block.

According to Grant Heffelfinger, there are four phases in Sandia's modeling history:

- Miscellaneous projects and researchers driven by a need to understand experiments characterized the pre-1988 era.
- 1988-1992 was characterized by algorithmic development, creating re-usable codes and methods for the high-end, funded by a variety of sources: the Grand Challenge, MICS (Mathematics, Information and Computer Sciences, Office of Computing, DOE), LDRD (Laboratory Directed Research & Development, tax on programs) and RF (Research Foundations—core funding from defense programs).
- 1992-1995 was characterized by collaborations with industry, funded through CRADAs. Materials modeling emerged as a strong theme.
- 1996-present (the ASCI era) has been characterized by the development of production codes, prototyping on stockpile problems, and a high-end focus.

The ASCI Program

ASCI (Accelerated Strategic Computing Initiative, http://www.sandia.gov/ASCI/) is an initiative of the Defense Programs Office at the U.S. Department of Energy in collaboration with LLNL (Lawrence Livermore National Lab), LANL (Los Alamos National Lab), and SNL (Sandia National Laboratories) to shift promptly from nuclear test-based methods to computation-based methods. The ASCI materials program at SNL has $5.5 million in funding, six projects, and over 35 PIs. The funding, however, is not quite adequate to expect to achieve fidelity in the materials simulation. The total ASCI program plans to provide simulation capabilities needed to predict the performance, safety, reliability, and manufacturing of weapons systems. Such simulations must be able to achieve high-resolution, three-dimensional calculations for full systems, using high-fidelity physical models.

MOLECULAR MODELING R&D ACTIVITIES

While the principal focus of the overall ASCI program is performance, the materials program spans all four aspects. Materials simulation and molecular-scale modeling are core to all the ASCI materials programs. There are six projects, all involving multiple length scales.

- Thermomechanical Fatigue (TMF) of Solder Joints
- Degradation of Organic Materials
- Corrosion
- Shock Response of PZT at the Grain-Scale
- Radiation Effects in Metals
- Solidification

Thermomechanical Fatigue (TMF) of Solder Joints

In the TMF program, the PARGRAIN code for grain growth and coarsening follows the microstructural evolution; it incorporates a strain-enhanced diffusion model and a decohesion model for micro-cracks and voids.

Degradation of Organic Materials

O-Ring aging or O-Ring lifetime prediction is a goal of the organic materials degradation project. The project utilizes a computationally efficient polymer builder code (SIGNATURE) which can build a 100% crosslinked network with up to 1.4 million atoms. Property prediction comes from MD/MC, TST, reactive MD, and QSPR. The results at these smaller length scales feed into longer-length-scale and longer-time-scale codes. Adhesive failure prediction is another goal in the organic materials degradation project. The objective is to develop the simulation capability to treat crack initiation at the molecular level.

Corrosion

The corrosion project focuses on sulfidation of Cu. Calculations at the atomistic scale use first principles DFT-based quantum calculations to obtain activation energies and mechanisms for point defect diffusion and to determine liquid-solid interfacial chemistry. The information at the atomic scale feeds codes for longer length scales.

Shock Response of PZT at the Grain-Scale

The goal of the PZT project is to determine the role of crystal and domain microstructure on the macroscopic electro-mechanical response of PZT (95% Zr, 5% Ti) with 2% Nb. The grain structure is produced by POROUS; LAGER converts the lattice-based grain structure to polyhedra. The calculations involve 1761 grains and 468,000 mesh elements, or ~250 elements per grain. The single-crystal material model incorporates data from first-principles DFT-based calculations for elastic constants. To determine the equilibrium phase diagram, first-principles "alchemical" perturbation methods are being developed. These methods use perturbation theory to convert Zr into Ti and therefore can work with small and ideal unit cells.

Radiation Effects in Metals

The molecular scale aspects of this project use the "embedded atom potential," validated with first principles codes.

Solidification

The unit processes at the atomistic scale are dislocations and defects; the microstructural scale is the grain scale. The microstructure determines the properties and the performance of a material. It is necessary to understand the development of the microstructure during processing and during operation in order to predict or optimize performance. Two dominant processes are interfacial energy minimization and bulk energy minimization. The suite of tools that Sandia researchers use includes MC-Potts, phase field, and front tracking, and they are exploring level-sets. "The best simulation is the one you only have to use once" as demonstrated by simulations of the grain aspect ratio of columnar growth as a function of pull speed. The simulations showed that the curve peaked at an optimal pull speed.

DISCUSSION

Sandia researchers identified several areas where breakthroughs could have large impacts:

- Design tools for micro-systems. Traditional tools are continuum-based; the physical basis for understanding is absent, since microsystems cannot rely on averaging. Furthermore, they have much greater surface/volume ratios. Lastly, they are hard to deal with experimentally.
- Biosystems
- Integration of software with experiments
- MD for very long time scales
- Chemometrics
- Sol-gel chemistry

Several other issues came up. First was the assertion that most materials codes are not mature enough to be used as black boxes. Second, there was concern that students in physical sciences are losing the training in analytical techniques. Students need to know that they are using codes right; hence, they need more education than just in the running of codes. They need some level of sophistication to determine a level of confidence in their results. Finally, there was an assertion that commercial software tends to be about five generations out of date.

CONCLUSIONS

The combination of having virtually no limitation on computational resources, a critical mass of researchers working at all length scales, and working together can produce some remarkable advances in codes and in applied problems that can be attacked. Despite the scale of this effort, there is concern that they are eating the seed corn, as there is not a commensurate investment in basic research developments in the missing fundamental science. Furthermore, nothing of this scale exists in industry, the coupling to commercial software is not optimal, and only a few select industries are even partially coupled to this effort.

Site: **Solutia Inc.**
3000 Old Chemstrand Rd
Cantonment, FL 32533
http://www.solutia.com/

Interview dates: 22 October 1999, 10 December 1999, 11 February 2000

Interviewer: P.R. Westmoreland

Interviewees: Genzo Tanaka, Research Specialist in Molecular Modeling
David L. Davidson, Fellow, dldavi@solutia.com

BACKGROUND

Solutia was formed in 1997 when the chemicals and materials portion of Monsanto (founded 1901) was spun off from Monsanto's agricultural chemicals and life sciences businesses. Company headquarters are in St. Louis, Missouri. Annual sales in 2000 were about $3 billion, and 11,000 people were employed worldwide.

The Pensacola (Cantonment) site is Solutia's largest facility, employing about 2,500 people for manufacturing nylon and nylon blends for fibers, carpet, tire yarn and specialty fibers applications. It also manufactures a variety of intermediate chemicals in addition to adipic acid and hexamethylene diamine (http://www.solutia.com/corporate/worldwide/pensacola.html).

MOLECULAR MODELING R&D ACTIVITIES

In 1995, the company began considering modeling to address polymer compatibility and to assist with the development of new catalyzed chemical processes. It purchased the pertinent MSI software and, in 1996, hired Dr. Genzo Tanaka in Pensacola. About this time, Dr. Freddy Zutterman in Belgium began applying molecular modeling to polymers and chemicals, primarily with MSI QSAR software on an Indigo2 computer.

As of early 2000, Solutia had two people who spend a significant part of their time in molecularly based modeling (Drs. Tanaka and Zutterman), two others who spend a smaller fraction of their time, and others who use it occasionally. Dr. David Davidson is responsible for a wide range of modeling, including computational fluid dynamics, optics, and mechanics. Dr. Tanaka is a graduate of the polymer chemistry department at Kyoto University, having postdoc-ed with Stockman at Dartmouth. His modeling has been mostly of polymers using MSI software. He also performs some PC-based analysis of data using neural nets.

An eight-processor SGI Origin 2000 was used at Pensacola for intensive computations, while PCs and SGI Indigo 2s were used for other processing. Solutia's experience with MSI has not been completely satisfactory, mainly because of maintenance issues and because new force fields are sold as extras, not supplied as updates.

Applications have fallen into four categories:

1. Screening to eliminate candidates and to search for new products or properties
2. Troubleshooting processes; e.g., responding to changes in manufacturing quality by using molecular modeling to test hypotheses about explanations
3. Property prediction; e.g., as glass transition temperature T_g, crystallinity, solvation properties
4. Catalysis; e.g., steric fit with respect to mass transport

DISCUSSION

Most of the work has been in-house and proprietary, although modest amounts of publication are possible. Calculations often have been fairly lengthy, such as a few weeks of computation for diffusivity. The calculations have proven useful and have revealed informative trends in data. However, there has been no single big success to point to, and Solutia researchers have recognized the value of developing one.

Two aspects were perceived as particular needs. First was the need for better force fields. They were using the Dreising force field most, but the different force fields have different pros and cons. Second, effective mesoscale modeling would be very valuable for predicting polymer microstructures and properties. The Mesodyne code of MSI seems promising.

They also believe it is important to share information about the nontechnical aspects of industrial molecular modeling. Although internal success stories are best, knowledge of external successes is also useful for establishing the value of the methods to internal clients. Likewise, it is useful to know different ways that molecular modeling works within other companies, how it may fail, and what are reasonable expectations.

POSTSCRIPT

Reduced sales and higher raw material costs led Solutia's leadership to reorganize the workforce. Consequently, while other modeling activities are alive and well (as of May 2001), molecular modeling as a specialized resource has been eliminated.

Internal analysis still led to the conclusion that when it becomes possible, Solutia should increase the resources devoted to molecular modeling and should modify the way it is applied. In particular, Solutia modelers recognized the need for a champion, somebody who would be both a fluent user of modeling tools as well as somebody who could sell it honestly and effectively to the enterprise.

In this context, it would be useful if molecular modeling tools were substantially easier to use. Solutia modelers have pressed CFD code vendors for some time to improve ease of use, including the addition of code to prevent engineers who are not experts in either CFD or fluid mechanics from getting into trouble—generating (and believing) erroneous results. If this approach were taken in molecular modeling software, it would allow companies without specialists to still get some value from molecular modeling.

Site: **University of Minnesota Supercomputing Institute**
1200 Washington Ave S.
Minneapolis, MN 55415
612-625-1818
http://www.msi.umn.edu

Date Visited: 6 December 1999

WTEC Attendees: E.B. Stechel (report author), P.T. Cummings, R. Miranda, P.R. Westmoreland

Hosts: Donald G. Truhlar, Computational Chemistry, Department of Chemistry, 612-624-7555, Fax 612-626-9390, truhlar@umn.edu
J. Ilja Siepmann, Computational Chemistry, Department of Chemistry, siepmann@chem.umn.edu
Darrin M. York, Computational Chemistry, Department of Chemistry, york@chem.umn.edu
James R. Chelikowsky, Computational Materials Science, Department of Chemical Engineering and Materials Science, jrc@msi.umn.edu

BACKGROUND

In 1981, the University of Minnesota was the first American university to acquire a supercomputer (a Cray-1B). The Supercomputing Institute was created in 1984 to provide leading-edge high-performance computing resources to the University of Minnesota's research community.

The Supercomputing Institute for Digital Simulation and Advanced Computation is an interdisciplinary research program spanning all colleges of the University of Minnesota. It covers all aspects of high-performance computing, scientific modeling, and simulation, as well as graphics, visualization, and high-performance network communications. In many cases, the research activities may involve efforts aimed at the design or evaluation of high-performance computing hardware, operating systems, networking, and general-purpose algorithms and software.

The Supercomputing Institute's resources are available to researchers at the University of Minnesota and other post-secondary educational institutions in the state of Minnesota. The Supercomputing Institute also organizes and hosts symposia, workshops, and seminars and coordinates educational and collaborative activities to promote supercomputing research, increase university-industry collaboration, and promote technology transfer.

The hardware resources have included a Cray-2, an ETA 10, a Cray X-MP, an IBM 3090, a Cray M90, and a Cray T3D. During 1997-98, the Supercomputing Institute provided access to a Cray T3E-900 and a Cray C90 system. The C90 had a total of twelve central processing units and contained 512 megawords of shared memory. The Cray T3E had 272 processors with a peak processing rate of 900 megaflops. In May 1998 the institute acquired a 256-processor (64 nodes with four 332 MHz 604e processors) IBM SP with 192 GB of memory (3 GB per node.) Before July 1998, management and resources of the Supercomputing Institute were provided by Network Computing Services (formerly Minnesota Supercomputer Center, Inc., or MSC Inc.). Since July 1998 the institute has been self-resourced.

In May 1999, the institute acquired a 160-processor IBM SP WinterHawk supercomputer with 80 GB of memory. In November 1999, the IBM SP resources were expanded to include a 68-processor IBM SP NightHawk supercomputer with 272 GB of memory and a 16-processor IBM SP Silvernode supercomputer with 12 GB of memory. Both WinterHawk and NightHawk supercomputers have 200 MHz Power3 processors. The Silvernode supercomputer has 200 MHz PowerPC 604e processors. The total amount of disk available to users of the IBM SP systems is currently 1.5 terabytes. The IBM supercomputer resources are available in coordination with the IBM Shared University Research (SUR) partnership.

In addition to the IBM SP resources, the Supercomputing Institute currently offers access to a 62-processor SGI Origin 2000 R12000 supercomputer, which has a clock speed of 300 MHz and a total of 50 GB of memory, and a 32-processor SGI Origin 2000 R10000 supercomputer with 8 GB of memory. The total amount of disk available to users of the SGI Origin 2000 systems is currently 640 GB.

MOLECULAR MODELING R&D ACTIVITIES

The seven principal faculty involved in research in computational chemistry and materials are Christopher J. Cramer, John S. Dahler, Jiali Gao, J. Ilja Siepmann, Donald G. Truhlar, and Darrin M. York of the Chemistry Department, and James R. Chelikowsky of the Department of Chemical Engineering and Materials Science.

The research efforts in the group of Chris Cramer range from the development and implementation of new theoretical models (e.g., the spin-annihilated density functional method, Class IV charge models, and solvation models) to the application of standard theoretical methods to answer chemical questions.

Research in John Dahler's group extends over several areas in nonequilibrium statistical mechanics and includes topics in atomic collision theory. Much effort is devoted to the kinetic theory of small molecule fluids with emphasis in (1) microscale fluid dynamics specific to the picosecond time scale and distances of a few angstroms; (2) transport, relaxation and light scattering by chemically reactive fluids; and (3) the dynamics of fast reactions and vibrational relaxation in liquids. Other topics of current concern are rubber elasticity, polymeric superionic conductors, the kinetic theory of suspensions and polymeric solutions, and the statistical thermodynamics of non-uniform fluids.

The Jiali Gao research group carries out theoretical and computational studies of structure, dynamics, and reactivity of biomacromolecules in solution. Of particular interest are enzymatic reactions and protein-substrate interactions. Additional work emphasizes combined quantum mechanical and molecular mechanical (QM/MM) techniques, and potential energy functions. Current projects include enzymatic reactions, organic and organometallic chemistry, and catalysis.

The goals in the research group of Ilja Siepmann include gaining microscopic-level insight into complex chemical systems, developing an understanding of the relationship between molecular architecture and macroscopic observables, and being able to make quantitative predictions of thermophysical properties. The algorithmic challenges include sampling of chain molecules, adiabatic sampling of nuclear and electronic degrees of freedom, connecting QM/MM/implicit solvation/mesoscale modeling. The challenges for force field development require balancing simplicity (to minimize computational cost) with transferability (to include state, composition and properties). The applications of interest include retention in gas-liquid and reversed-phase liquid chromatography, supercritical fluid extraction systems, surfactantcy, and organic thin films (SAMs and Langmuir monolayers).

The research group of Don Truhlar does theoretical and computational studies of chemical reaction dynamics, potential energy surfaces, and molecular solvation. The tools are quantum mechanics, classical and semiclassical mechanics, and molecular modeling. Additional work focuses on molecular energy levels and thermochemistry. Application areas include quantum photochemistry (electronically non-adiabatic processes), combustion, biomolecules, catalysis, high-energy species, and atmospheric and environmental chemistry.

The focus of Darrin York's research group is to develop and apply new theoretical methods to study biological macromolecules in solution. Methodological developments in the lab include new quantum mechanical (electronic structure) methods, in addition to novel molecular simulation techniques and new generation force fields that include many-body effects. At the interface of these methods are hybrid QM/MM methods for studying detailed chemical reactions of biological macromolecules in solution. The application focus consists of challenging problems involving ionic biosystems including metal ion catalysis, phosphate hydrolysis reactions, and protein-nucleotide interactions that are beyond the scope of conventional empirical force fields.

Jim Chelikowsky's group focuses on developing new algorithms and methods to predict properties of real materials. Specific interests include liquid semiconductors, pressure-induced amorphization in silica and related materials, semiconductor clusters, and defects on surfaces.

DISCUSSION

A number of software programs have been developed in the Computational Chemistry Group and distributed from the University of Minnesota, including ABCRATE, MOPAC 5.09mn, AMSOL, MORATE, AMSOLRATE, MULTILEVEL, CRATE, OMNISOL, GAMESOL, POLYRATE, GAMESOLRATE, RMPROP, GAUSSRATE, and TINKERATE. These are available through the group's Web site.

The center's proposed set of computing resources for 2000-2001 reflects the general move away from Cray.

In reflection upon possible breakthroughs that could have significant effect, the hosts identified four areas:

- Integration of a hierarchy of levels of model accuracy and experiment
- Good methodologies for parameterizing systems with constraints from experiments, as one cannot attack complex systems entirely from first principles
- Force field developments that optimally balance simplicity and transferability
- Connecting the length scales

CONCLUSIONS

The panel particularly wished to visit the Supercomputing Institute because of its parallels with the Institute for Molecular Science in Japan (see site report). The institute's computational chemistry and computational materials groups are making serious attempts to widely distribute their theoretical and computational advances in the form of freely available software. Some faculty members have prior industrial experience and several have active collaborations with industry. Contacts and the matching of interests across the academic/industrial sectors typically occur at conferences. Examples of interactions with industry include collaborations with Phillips Petroleum, Kodak, Entropic Systems Inc., and Rohm and Haas. Some interactions are funded by GOALI and SBIR grants.

REFERENCES

For a more complete list of references for D.G. Truhlar, see http://comp.chem.umn.edu/truhlar/

Alhambra, C., J. Gao, J.C. Corchado, J. Villà, and D.G. Truhlar. 1999. Quantum mechanical dynamical effects in an enzyme-catalyzed proton-transfer reaction. *Journal of the American Chemical Society* 121, 2253-2258.

Chuang, Y.-Y., and D.G. Truhlar. 1999. Geometry optimization with an infinite basis set. *Journal of Physical Chemistry A* 103, 651-652.

Chuang, Y.-Y., and D.G. Truhlar. 1999. Nonequilibrium solvation effects for a polyatomic reaction in solution. *Journal of the American Chemical Society* 121, 10157-10167.

Chuang, Y.-Y., J.C. Corchado, and D.G. Truhlar. 1999. Mapped interpolation scheme for single-point energy corrections in reaction rate calculations and a critical evaluation of dual-level reaction-path dynamics methods. *Journal of Physical Chemistry A* 103, 1140-1149.

Chuang, Y.-Y., M.L. Radhakrishnan, P.L. Fast, C.J. Cramer, and D.G. Truhlar. 1999. Direct dynamics for free radical kinetics in solution: Solvent effect on the rate constant for the reaction of methanol with atomic hydrogen. *Journal of Physical Chemistry* 103, 4893-4909.

Cramer, C.J., and D.G. Truhlar. Implicit solvation models: equilibria, structure, spectra, and dynamics. *Chemical Reviews* 99, 2161-2200.

Espinosa-García, J., J.C. Corchado, and D.G. Truhlar. 1997. The importance of quantum effects for C-H bond activation reactions. *Journal of the American Chemical Society* 119, 9891-9896.

Fast, P.L., and D.G. Truhlar. 1999. Simple approximation of core-correlation effects on binding energies. *Journal of Physical Chemistry A* 103, 3802-3803.

Fast, P.L., J. Corchado, M.L. Sanchez, and D.G. Truhlar. 1999. Optimized parameters for scaling correlation energy. *Journal of Physical Chemistry A* 103, 3139-3143.

Fast, P.L., J.C. Corchado, M.L. Sánchez, and D.G. Truhlar. 1999. Multi-coefficient correlation method for quantum chemistry. *Journal of Physical Chemistry A* 103, 5129-5136.

Fast, P.L., M.L. Sánchez, and D.G. Truhlar. 1999. Infinite basis limits in electronic structure theory. *Journal of Chemical Physics* 111, 2921-2926.

Fast, P.L., M.L. Sánchez, and D.G. Truhlar. 1999. Multi-coefficient Gaussian-3 method for calculating potential energy surfaces. *Chemical Physics Letters* 306, 407-410.

Fast, P.L., M.L. Sánchez, J.C. Corchado, and D.G. Truhlar. 1999. The Gaussian-2 method with proper dissociation, improved accuracy, and less cost. *Journal of Chemical Physics* 110, 11679-11681.

Friedman, R.S., T.C. Allison, and D.G. Truhlar. 1999. Exciplex funnel resonances in chemical reaction dynamics: The nonadiabatic tunneling case associated with an avoided crossing at a saddle point. *Physical Chemistry Chemical Physics 1*, 1237-1247 (special issue on chemical reaction theory).

Hack, M.D., A.W. Jasper, Y.L. Volobuev, D.W. Schwenke, and D.G. Truhlar. 1999. Quantum mechanical and quasiclassical trajectory surface hopping studies of the electronically nonadiabatic predissociation of the $\tilde{A}$ state of NaH_2. *Journal of Physical Chemistry A* 103, 6309-6326.

Hack, M.D., and D.G. Truhlar. 1999. Analytic potential energy surfaces and their couplings for the electronically nonadiabatic chemical processes $Na(^3p) + H_2$ -> $Na(^3s) + H_2$ and $Na(^3p)$ + H2 -> NaH + H. *Journal of Chemical Physics* 110, 4315-4337.

Li, J., C.J. Cramer, and D.G. Truhlar. 1999. Application of a universal solvation model to nucleic acid bases. Comparison of semiempirical molecular orbital theory, ab initio Hartree-Fock theory, and density functional theory. *Biophysical Chemistry* 78, 147-155 (special issue on implicit solvent representation in biomolecular chemistry).

Li, J., J. Xing, C.J. Cramer, and D.G. Truhlar. 1999. Accurate dipole moments from hartree-fock calculations by means of class IV charges. *Journal of Chemical Physics* 111, 885-892.

Li, J., T. Zhu, G.D. Hawkins, P. Winget, D.A. Liotard, C.J. Cramer, and D.G. Truhlar. 1999. Extension of the platform of applicability of the SM5.42R universal solvation model. *Theoretical Chemistry Accounts* 103, 9-63.

Roberto-Neto, O., F.B. C. Machado, and D.G. Truhlar. 1999. Energetic and structural features of the CH4 + $O(^3P)$ -> CH_3 + OH abstraction reaction: Does perturbation theory from a multiconfiguration reference state (finally) provide a balanced treatment of transition states? *Journal of Chemical Physics* 111, 10046-10052.

Srinivasan, J., T.C. Allison, D.W. Schwenke, and D.G. Truhlar. 1999. Transition state resonances in the reaction Cl + H_2 -> HCl + H. *Journal of Physical Chemistry A* 103, 1487-1503.

Topaler, M.S., P. Piecuch, and D.G. Truhlar. 1999. Infrared absorption line strengths of the Na...FH van der Waals molecule. *Journal of Chemical Physics* 110, 5634-5638.

Villà, J., J.C. Corchado, A. González-Lafont, J.M. Lluch, and D.G. Truhlar. 1999. Variational transition state theory with optimized orientation of the dividing surface and semiclassical tunneling calculations for deuterium and muonium kinetic isotope effects in the free radical association reaction H + C2H4 -> C2H5. *Journal of Physical Chemistry A* 103, 5061-5074.

Volobuev, Y.L., M.D. Hack, and D.G. Truhlar. 1999. Are semiclassical methods accurate for electronically nonadiabatic transitions between weakly coupled potential energy surfaces? *Journal of Physical Chemistry A* 103, 6225-6233.

Zeiri, Y., G. Katz, R. Kosloff, M.S. Topaler, D.G. Truhlar, and J.C. Polanyi. 1999. Quantum mechanism in the photodissociation of NaFH complex: A challenge to semiclassical analysis. *Chemical Physics Letters* 300, 523-528.

Zhu, T., J. Li, D.A. Liotard, C.J. Cramer, and D.G. Truhlar. 1999. Analytical energy gradients of a self-consistent reaction-field solvation model based on CM2 charges. *Journal of Chemical Physics* 110, 5503-5513.

For a more complete list of references for J.I. Siepmann, see: http://siepmann6.chem.umn.edu/~group/publications.html

Martin, M.G., J.I. Siepmann, and M.R. Schure. 1999. Simulating retention in gas-liquid chromatography. *J. Phys. Chem. B* 103, 11191-11195.

Siepmann, J.I. 1999. An introduction to the Monte Carlo method for particle simulations. *Adv. Chem. Phys.* 105, 1-12.

Siepmann, J.I. 1999. Monte Carlo methods for simulating phase equilibria of complex fluids. *Adv. Chem. Phys.* 105, 443-460.

Siepmann, J.I., and I.R. McDonald. 1993. Monte Carlo simulation of the mechanical relaxation of a self-assembled monolayer. *Phys. Rev. Lett.* 70, 453-456.

Siepmann, J.I., and I.R. McDonald. 1998. Simulations of self-assembled monolayers of thiols on gold. *Thin Films* 24, 205-226.

Siepmann, J.I., S. Karaborni, and B. Smit. 1993. Simulating the critical properties of complex fluids. *Nature* 365, 330-332.

Smit, B., and J.I. Siepmann. 1994. Simulating the adsorption of alkanes in zeolites. *Science* 264, 1118-1120.

For a more complete list of references for Darrin York, see http://pinot-noir.chem.umn.edu/~york/

York, D.M., T.-S. Lee, and W. Yang. 1996. Parameterization and efficient implementation of a solvent model for linear-scaling semiempirical quantum mechanical calculations of biological macromolecules. *Chemical Physics Letters* 263 297-304.

York, D.M., T.-S. Lee, and W. Yang. 1996. Quantum mechanical study of aqueous polarization effects on biological macromolecules. *Journal of the American Chemical Society* 118 10940-10941.

York, D.M., T.-S. Lee, and W. Yang. 1998. Quantum mechanical treatment of biological macromolecules in solution using linear-scaling electronic structure methods. *Physical Review Letters* 80 5011-5014.

York, D.M., W. Yang, H. Lee, T. Darden, L. Pedersen. 1995. Toward the accurate modeling of DNA: The importance of long-range electrostatics. *J. Am. Chem. Soc. (communication)* 117, 5001-5002.

For a more complete list of J.R. Chelikowsky references see: http://jrc.cems.umn.edu/

Chelikowsky, J.R., and N. Binggeli. 1998. Modeling the properties of quartz with clusters. *Solid State Communications* 107, 527.

Godlevsky, V.V., and J.R. Chelikowsky. 1998. Ab initio molecular dynamics simulations of liquid GaAs. *J. Chem. Phys.* 109, 7312 (1998).

Godlevsky, V.V., J.J. Derby and J.R. Chelikowsky. 1998. Ab initio molecular dynamics simulation of liquid CdTe and GaAs: Semiconducting versus metallic behavior. *Physical Review Letters* 81, 4959.

Jay, L.O., H. Kim, Y. Saad and J.R. Chelikowsky. 1999. Electronic structure calculations using plane wave codes without diagonalization. *Computer Physics Communications* 118, 21.

Ogut, S., and J.R. Chelikowsky. 1999. Large pairing Jahn-Teller distortions around divacancies in crystalline silicon. *Phys. Rev. Lett.* 83, 3852.

Vasiliev, I., S. Ogut, and J.R. Chelikowsky 1999. Ab initio excitation spectra and collective electronic response in atoms and clusters. *Physical Review Letters* 82, 1919.

Vasiliev, I., S. Ogut, and J.R. Chelikowsky. 1999. Ab initio absorption spectra of gallium arsenide clusters. *Phys. Rev. B* 60, R8477.

Vasiliev, I., S. Ogut, and J.R. Chelikowsky. 1999. Ab initio excitation spectra and collective electronic response in atoms and clusters. *Physical Review Letters* 82, 1919.

Wentzcovitch, R.M., C. da Silva, J.R. Chelikowsky, and N. Binggeli. 1998. A new phase and pressure induced amorphization in silica. *Phys. Rev. Lett.* 80, 2149.

APPENDIX C. OBSERVATIONS ON FUNDING ESTIMATES FOR MOLECULAR MODELING

Emile Ettedgui, RAND Corporation

INTRODUCTION

On December 8, 1999, Geoffrey Prentice of NSF discussed his interest in obtaining information on federal funding devoted to molecular modeling. Based on the discussion, the funding information should be broad and cover all agencies, as applicable. While the theme of the research projects should be molecular modeling, this term should include the different meanings that it may take by discipline.

Following this discussion, I carried out a RaDiUS[4] search on molecular modeling and related terms. The results are included in this memo, along with observations resulting from this effort.

METHODOLOGY

The search was carried out for fiscal year (FY) 1998. This is the latest year for which complete information is available. During the course of this search, a number of search terms were generated. These were tried in turn, while being careful to eliminate double counting of entries by specifying that previously matched terms be excluded. Thus, if the search began with term A and was expanded to term B, the second iteration selected matches for term B that did not also match term A. This approach was continued as new terms were added to the search.

Following each iteration of the search, the results were scanned rapidly to determine whether they were relevant to the project. If the search appeared successful, the results were printed and each abstract was read for relevance to the search. If the search did not appear successful, the search term was refined or changed to improve the relevance of the result.

RESULTS AND DISCUSSION

Based on the iterative search, the search terms used are as follows. The terms are listed in the order that they were used for the search. Each entry represents the actual term used. An explanation of the term used appears in italics.

- molecul% model% — *molecular modeling and related terms*
- ab initio — *ab initio*
- first principl% — *first principle and first principles*
- atom% orbita% — *atomic orbital and related terms*
- molec% orbita% — *molecular orbital and related terms*
- molec% dyn% — *molecular dynamics and related terms*
- hartr% — *Hartree-Fock and similar terms*

---

[4] Editor's note: RaDiUS is a union database of U.S. government research and development funding activities. See http://www.rand.org/scitech/radius/ for more information.

- oppenheim% — *Born-Oppenheimer and similar terms*
- (molecul%) AND (monte carlo) AND NOT (gene%) — *molecular and related terms and Monte Carlo, but excluding genetics and related terms*

The last term given illustrates the kind of modifications that were necessary during the course of the data search. Monte Carlo simulation includes a very large number of terms that are not related to molecular dynamics. Even restricting the search of Monte Carlo simulations to ones including the term molecule and related terms, a number of terms related to molecular genetics appeared among the matches. These were not related to molecular modeling as such and were therefore excluded.

Looking at the number of terms matched for each search term suggests that the number of matches is decreasing as new terms are introduced. The number of matches given corresponds to the number of records retrieved and does not take into account any further screening that resulted from reading the abstracts. This is summarized as follows, where the number of matches is shown for each new term in the search:

- molecul% model% 335
- ab initio 175
- first principl% 168
- atom% orbita% 2
- molec% orbita% 36
- molec% dyn% 612
- hartr% 18
- oppenheim% 10
- (molecul%) AND (monte carlo) AND NOT (gene%) 33

The total number of matches was 1389. As each new search term was introduced, the previously matched terms were excluded from the search. As a result, if the search terms had been introduced in a different order, the number of matches for a given term may have been different. A search including all terms would also have been possible. The reason that the search was carried out in an iterative fashion by excluding previous terms was that the additional terms generated could be easily separated from previous matches and reviewed individually.

The results presented may not correspond to an exhaustive search. This is because additional search terms may uncover further matches within RaDiUS. Even so, it is felt that the information gathered at this point may provide a good estimate of funding efforts for molecular modeling. This is because four of the terms contribute 95% of the matches and because of the declining contributions from additional search terms.

After a preliminary screening of the abstracts to determine the suitability of the search terms employed, the abstracts were read individually and scored. This was done because the RaDiUS search engine matches terms rather than meaning. An abstract that refers to a given search term will thus be selected, even if the proposed work does not deal with this subject. Scores between 0 and 4 were assigned to each abstract. A project unrelated to molecular modeling was scored 0 and one that deals with molecular modeling was scored 4. Intermediate scores were used when descriptions did not provide an unequivocal answer.

The level of funding associated with research on molecular modeling was then calculated by counting all projects with a score of 4. The budgets of projects with a score of 3 were counted at 75%, since it is likely that these projects deal with molecular modeling, although one cannot be certain from the project description. Using a similar approach, projects with a score of 2 were counted at 50% and those with a score of 1 were counted at 25%. Projects with a score of 0 were not counted.

The relative contribution of each project was then multiplied by its average annual funding available in the RaDiUS database. In the case of the Department of Defense projects, the annual funding was estimated by dividing the program obligation by the number of projects that it included. Based on this approach, the total

budget devoted to molecular modeling was $158,122,000. The budget devoted to molecular modeling was also estimated without scaling the contributions of the various abstracts, in which case it equaled $207,588,000.

Table C.1 below summarizes funding and project information for the various government agencies. The first column provides the name of the agency. The second column shows how often budget information is available within a given agency. The first number given in the second column indicates the number of projects for which funding information is available. The number in parentheses in the second column shows how many projects were funded by a given agency. The third column gives the annual funding level that results without taking the score of each project into account. In the fourth column, the funding is provided based on the score obtained for each abstract. This is the figure that is used to compute the total funding devoted to molecular modeling. In all cases, funding levels are provided in thousands of dollars.

With the exception of Department of Defense projects, the vast majority of the project descriptions contain cost information. Some projects in other agencies did not contain cost information either. However, these represent a minority contribution. This is based on the number of projects dealing with molecular modeling and related work.

Estimates for Department of Defense efforts in molecular modeling are calculated differently from those in other agencies. This is because the Department of Defense does not provide information on the funding of individual research efforts. RaDiUS provides funding information within agencies, however. By dividing funding for research by the number of projects within agencies, it is possible to estimate the level of funding of molecular modeling that originates within the Department of Defense.

Table C.1
RaDiUS Estimates of FY 1998 U.S. Government Funding for Molecular Modeling ($ thousands)

Government Agency	Funding Figures Available	Average Annual Funding (AAF)*	Scaled AAF*
Department of Agriculture	4 (10)	547	547
Department of Commerce	6 (6)	11,419	6,172
Department of Defense	0 (203)	16,248	9,010
Department of Energy	54 (55)	25,518	19,446
Department of Health and Human Services	472 (520)	116,016	91,090
Department of Veterans Affairs	24 (24)	1,167	1,060
Environmental Protection Agency	4 (4)	1,055	643
National Aeronautics & Space Administration	6 (22)	278	273
National Science Foundation	507 (509)	35,340	29,878

* Due to rounding, the total may not match exactly the one given in the text.

For comparison, the results of this search are presented along with information provided by the National Science Foundation. The first column in Table C.2 lists the various directorates with the NSF in alphabetical order. In the second column are listed NSF estimates for resources devoted to molecular modeling within the various directorates. The third column shows the results obtained from the RaDiUS search based on the methodology described above. Funding levels are provided in thousands of dollars.

It is also possible to categorize recipients of funding for molecular dynamics[5] based on their affiliation (see Table C.3). In this case, the recipients are split into six groups. These are: educational institutions, the federal

[5] Editor's note: molecular modeling

government, industry, non-profit non-educational institutions, state and local governments, and non-specified institutions. The summary appears below. Again the figures are presented in thousands of dollars.

Table C.2
Comparison of Internal NSF Funding Estimates and RaDiUS Estimates, FY 1998 ($ thousands)

NSF Directorate	NSF estimate	AAF	Scaled AAF
Biological Sciences	> 4,000	4,288	3,920
Computer &Information Science & Engineering	?	2,185	1,422
Education &Human Resources	?	616	175
Engineering	> 3,200	4,534	3,663
Geosciences	?	923	779
Mathematical & Physical Sciences	> 15,000	22,291	19,770
Social, Behavioral & Economic Sciences	?	503	151

Table C.3
Recipients of Federal Funding for Molecular Modeling Research ($ thousands)

	AAF	Scaled AAF
Educational	117,870	94,364
Federal government	48,103	34,704
Industry or business	15,105	8,129
Non-profit, non-educational	20,949	16,007
State or local government	3,207	3,117
Not specified	2,355	1,802

APPENDIX D. GLOSSARY

Ab initio methods—Computational approaches that attempt to solve the molecular electronic Schroedinger equation without using any empirical parameters.

Active analog approach—A method, developed by Marshall and co-workers in the late 1960s and early 1970s, that attempts to ascertain the structural, electrostatic and hydrophobic features of the active molecules by superimposing active and inactive analogs.

ADF—Amsterdam Density Functional program, a computer program for carrying out density functional quantum mechanical calculations; developed at the University of Amsterdam and University of Calgary (http://www.ccwp.ac.uk/ccwp/adf.html) and distributed by MSI.

ADME properties—Adsorption, distribution, metabolism and excretion. Properties of molecules crucial to their *in vivo* biological activities which are not represented if one only considers the non-covalent interactions of ligand and receptor.

AFM—Atomic force microscopy.

AMBER—Assisted model building with energy refinement. Force fields and a computer program for carrying out molecular mechanics and dynamics of organic and biological molecules developed at the University of California, San Francisco, by Peter Kollman and co-workers. See http://www.amber.ucsf.edu/amber/amber.html.

ASCI—Accelerated Strategic Computing Initiative. An initiative to carry out simulations rather than field tests to maintain the viability of US nuclear weapons. The program resulted in creation of the four fastest supercomputers in the world (http://www.top500.org/, November 2000) with peak speeds up to 12 teraflops: ASCI White (http://www.llnl.gov/asci/), ASCI Red (http://www.sandia.gov/ASCI/Red/), ASCI Blue Pacific (http://www.llnl.gov/asci/platforms/bluepac/), and ASCI Blue Mountain (http://www.lanl.gov/projects/asci/bluemtn/bluemtn.html). See also http://www.asci.doe.gov/.

ATP—Advanced Technology Program of the National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce. ATP is intended to aid higher-risk research and development by U.S. companies, co-funding work with industry that may involve government and academic partners. (See also adenosine triphosphate.)

AUTODOCK—A computer program for docking of ligands into macromolecular sites, developed at Scripps Research Institute.

Basis set—A parameterized set of functions representing the wavefunctions of atomic orbitals, used to construct molecular orbitals in electronic-structure calculations.

Bead-spring models—Simplified models to represent polymers or proteins that do not use all atom representations.

Bioinformatics—The field whose focus is to represent and manipulate the genetic information coming from the explosion of genomics.

Born-Oppenheimer approximation—The approximation that allows one to solve the electronic Schroedinger equation with fixed nuclei. It is a good approximation for ground states and probably for many excited states, except those with curve crossings.

CAChe—A software suite marketed by Fujitsu, originally developed by Oxford Molecular Group. Useful for semi-empirical molecular-orbital calculations, visualization, properties, and information management. See http://www.cachesoftware.com/.

CAE—Computer-aided engineering. Use of computational methods in practical aspects of engineering design.

Canonical ensemble—A statistical mechanical ensemble that represents the system in terms of constant volume, temperature and number of particles.

Car-Parrinello molecular dynamics—An approach to solving electronic quantum mechanical equations using classical dynamics for the nuclei and fictitious masses for the electrons, developed by R. Car and M. Parrinello.

Carbon nanotubes—Large clusters of carbon atoms in the shape of a tube with nanometer dimensions.

CASTEP—A software package for *ab initio* solid-state calculations developed at Cambridge, England, marketed by MSI / Accelrys .

Catalysis—Lowering the free energy of the transition state for a chemical reaction using, for example, an enzyme or inorganic molecule.

Ceramics—A set of materials with particular solid-state properties.

CFD—Computational fluid dynamics. An approach to solving the differential equations for fluid flow using computational techniques.

CFS—Computing for Science Ltd. . A software company founded in 1992 by a consortium led by Martyn Guest (Daresbury), J.H. van Lenthe (Utrecht), J. Kendrick (ICI), and K. Schoffel (Norsk Hydro) that markets the GAMESS-UK *ab initio* code. See http://www.dl.ac.uk/CFS/.

CHARMM—Chemistry at Harvard molecular mechanics. A computer program for molecular mechanics and dynamics developed by Karplus and co-workers, distributed by Harvard University (http://yuri.harvard.edu/) and by MSI.

Chemical informatics or cheminformatics—The discipline of organizing and analyzing information about the properties of molecules, a term currently most used in drug development and discovery but not restricted to that field.

Cluster—An approximation for carrying out computational studies on surfaces where one only includes a shell of atoms/molecules surrounding the interesting region.

Combinatorial chemistry—The use of combinatorial methods to synthesize many molecules based on a common scaffold and many possible substituents coming off that scaffold.

Computational chemistry—Typically, calculation of molecular structure, properties, or reactivity using quantum mechanics (computational quantum chemistry); in a more general sense, any use of computers applied to chemistry problems.

COMFA—Comparative mean field analysis. A 3D QSAR approach that attempts to analyze the space around a set of molecules to determine those criteria that determine activity or inactivity in binding or some other biological effect. This approach is a quantitation of the active analog approach.

Configuration interaction (CI)—An approach to include correlation effects in a computation of molecular electronic structure using a multi-determinant wave function.

Continuum Configuration-Bias Monte Carlo—A method to carry out Monte Carlo simulations for dense liquids that enables reasonable acceptance criteria, which is not the case with standard MC. It involves placing the molecules into the liquid phase one monomer at a time and choosing the location of each new monomer in an optimum fashion.

CRADA—Cooperative research and development agreement. Research and development activity by Federal and industrial researchers working together, typically with negotiated confidentiality of results.

DACAPO—A software package for *ab initio* solid-state calculations developed in Denmark by Nørskov and co-workers (see CAMP site report in Appendix B.1).

DARPA—Defense Advanced Research Projects Agency. A central research and development organization for the U.S. Department of Defense, conducting and sponsoring basic and applied R&D. See http://www.darpa.mil/.

Density functional methods (DFT)—Methods to solve for the electronic structure of molecules that make use of the fact that the solution is a function only of the electron density. These methods are "almost *ab initio*," but the exchange correlation functional that is part of the solution is partly empirical and/or based on model systems.

Discover—A computer program for molecular mechanics and dynamics developed at Biosym, now part of Molecular Simulations, Inc. / Accelrys . See http://www.msi.com/.

Dislocation dynamics in metals—The way metals deform under stress and the way these deformations are propagated in time.

Distance geometry—An approach that uses geometry and upper and lower bounds on distances to derive three dimensional structures for complex systems.

Divide-and-conquer methods—Methods which attempt to solve for the electronic structure of large molecules using approximations to make the algorithm approach linear dependence on the size of the system.

DOCK—A computer program for docking molecules together developed at the University of California, San Francisco.

Docking—Methods to predict how two molecules (small-large or large-large) optimally fit together in a non-covalent complex.

DoD—U.S. Department of Defense (http://www.defenselink.mil/).

EA or EH—Explicit-atom model for molecular simulations, in which all atoms are included explicitly (vs. UA).

Effective core potential—An approach to put the inner shells of electrons in an effective potential while solving for the electronic structure of the valence electrons of the system; may include relativistic effects that may be important for electrons near large nuclei.

Embedded cluster or embedded atom methods—Approaches to calculate the properties of a cluster of molecules on a surface by including an approximation for the rest of the environment not included in the explicit calculation.

EMSL—William R. Wiley Environmental Molecular Science Laboratory. A laboratory of DOE's Pacific Northwest National Laboratory (Richland, Washington) focused on research in computational and molecular sciences to solve environmental restoration problems. See http://www.emsl.pnl.gov.

EPSRC—Engineering and Physical Sciences Research Council. Largest of the seven UK Research Councils. See http://www.epsrc.ac.uk/.

EXAFS—Extended X-ray absorption fine structure. An experimental approach to derive structural properties of ligands surrounding a metal.

Expert systems—A computer approach using artificial intelligence to analyze the properties of complex systems.

Force fields—Analytical expressions that describe the potential energy of systems as a function of the positions of the nuclei.

FLEXX—A docking program developed by Rarey, Lengauer, et al., in an academic group in Germany; distributed by Tripos, Inc.

Flory chi parameter or Flory-Huggins interaction parameter χ_1—Energy required to remove a mole of solvent molecules from the bulk solvent and dissolve it into a polymer.

Fast multipole method (FMM)—An $O(N)$ algorithm developed by Greengard and Rokhlin to compute electrostatic energy in MD simulation, subdividing physical space into cells to compute short-range interactions within the cells and long-range interactions between them.

FTIR—Fourier transform infrared spectroscopy. Using Fourier-transform techniques to measure the infrared spectra of a molecular system.

G1, G2, or G3 data sets and methods—Certain test sets of molecules with well-characterized properties, used with their corresponding prescribed sets of quantum-chemistry calculations to calibrate empirical corrections and represent what level of error this kind of model can achieve relative to experiment; due to Pople and co-workers.

GAMESS-UK, GAMESS-US, PC GAMESS—General-purpose *ab initio* molecular electronic structure programs based on the original program GAMESS developed in 1980 at NRCC of Lawrence Berkeley Laboratory by M. Dupuis, D. Spangler, and J.J. Wendoloski. See:
http://www.dl.ac.uk/TCSC/QuantumChem/Codes/GAMESS-UK/
http://www.msg.ameslab.gov/GAMESS/GAMESS.html
http://classic.chem.msu.su/gran/gamess/.

Gaussian XX—A computer program for various methods of computational quantum chemistry. It was originally developed in the laboratory of John Pople (GAUSSIAN 70) and has been distributed since 1987 by a commercial company (Gaussian, Inc. , Pittsburgh PA and North Haven CT; http://www.gaussian.com/) that has developed subsequent versions of this software. XX approximately represents the year of appearance.

Genomics—Computationally analyzing and manipulating information about chemically encoded genetic material (DNA or the genome).

Gibbs-ensemble Monte Carlo—A particular version of Monte Carlo molecular simulations that uses different simulation boxes and exchanges particles between them at constant temperature, pressure, and chemical potential. The usual purpose is to calculate phase equilibria by establishing the relative free energies of a given molecule in two very different phases.

Grand canonical ensemble—A statistical mechanical ensemble with constant temperature, volume and chemical potential, in which the number of particles can change to keep the chemical potential constant.

Grid computing—Computing across many platforms, using features of different hardware and software platforms.

GROMOS—Groningen molecular simulations. A software package for molecular mechanics and dynamics developed by W.F. van Gunsteren at Harvard, Groningen and ETH Zürich, and maintained at ETH Zürich. See http://www.rereth.ethz.ch/chem/physikalische/vangunst/pj.08.html.

Hansch approach—Using linear regression to enable QSAR (quantitative structure-activity relationships) to be derived to relate physical properties to biological activity.

Hartree-Fock theory—Electronic structure calculations that use a single-determinant wavefunction and thus do not include any correlation energy. Such calculations carried out at the *ab initio* level with a very large basis set approach an energy that is called the Hartree-Fock limit.

Heterogeneous catalysis—Catalysis at interfaces, not in pure media.

Histogram reweighting—An approach to computing a density distribution for a molecular system in a grand canonical ensemble and inferring from this the densities of the coexisting phases, as well as estimating density distributions at different chemical potentials and temperatures from one simulation.

Homogeneous catalysis—Catalysis by molecules, usually organometallic, in homogeneous media.

HPC—High-performance computing. Applications of supercomputers.

HREELS—High-resolution electron energy loss spectroscopy. A surface spectroscopy method used to examine vibrations of molecules on surfaces.

HRTEM—High-resolution transmission electron microscopy.

Hyperchem—Software package developed at Hypercube, Inc. (Gainesville, FL) that includes both molecular mechanical and quantum mechanical capabilities. See http://www.hyper.com/.

Informatics—Computational analysis of information.

Insight—A computer program for molecular visualization developed at the company Biosym, which is now part of MSI.

Interatomic potentials—Analytical potential functions for atomic or molecular interactions as a function of nuclear position.

Interfaces—Used either in a software context as a program to connect pieces of software or in a chemical context to describe the interface between two phases (e.g., solid and liquid). Often the location for heterogeneous chemical catalysis.

Intrinsic reaction coordinate—The least-energy pathway from reaction to product, used in quantum chemical software packages to study the potential energy surfaces for chemical reactions.

I/O—Input/output or input/output speed.

INITR (or IT^2)—Information Technology Initiatives proposed by the Clinton administration and passed by Congress, enabling funding of computer science and computational science as well as hardware.

ISV—Independent software vendor.

JAGUAR—A software package for quantum mechanical electronic structure calculations developed by the company Schrödinger, Inc. (http://www.schrodinger.com/).

Kinetic Monte Carlo—A method to model reaction kinetics within a molecular simulation framework by using stochastic transitions.

LBL—Lawrence Berkeley (National) Laboratory, U.S. DOE national laboratory in Berkeley, CA. See http://www.lbl.gov/.

Library development—The development of classes of chemical compounds as part of a library, an alternative phrase might be "combinatorial library development."

Ligand—A chemical component that is bonded to another.

LLNL—Lawrence Livermore National Laboratory, U.S. DOE national laboratory in Livermore, CA. See http://www.llnl.gov/.

LNG—Liquefied natural gas.

MIDAS—A software package for graphical visualization developed at the University of California, San Francisco.

Molecular dynamics—A method to carry out a classical trajectory of a molecular system, numerically solving Newton's law of motion and using a particular force field for the potential energy of the system.

Molecular mechanics—The use of an analytical force field to describe the potential energy of the system as a function of nuclear position. Molecular mechanics can be thought of as a composite of molecular dynamics and Monte Carlo calculations, although in its usage it is often synonymous with "energy minimization," which is molecular dynamics at 0 K.

Molecular modeling or molecularly based modeling—The use of physical models or computational models to simulate the properties of a molecular system.

Molecular simulations—Usually molecular dynamics or Monte Carlo calculations, which involve deriving a trajectory or state for a molecular system that can be related to a statistical mechanical ensemble.

Molecular Simulations, Inc.—A company based in San Diego, CA and in Cambridge, UK (a subsidiary of Pharmacopeia, Inc.—see site reports) that sells computational chemistry and informatics software. It has foci in both biological and materials-science modeling arenas. The latter area includes modeling of non-biological chemistry. Along with other software companies within Pharmacopeia, it was reformed as Accelrys on June 1, 2001. See http://www.msi.com/.

Monte Carlo simulations—An approach to molecular simulations in which random moves of the molecules are made and accepted with a Boltzmann criterion.

Moore's law—The hypothetical increase in computer speed as a function of time, typically a factor of 2 every 18 months; hypothesized in 1965 by Gordon Moore as doubling the number of transistors on a integrated circuit every year, based on 1 in 1959 with intermediate steps to about 64 in 1965 (Moore, G.E. 1965. *Electronics*. 38:8. April 19.) In 1997 it was revised to doublng of performance every 18-24 months (Moore, G.E. 1965. Intel Deeloper's Forum. San Francisco, CA. September 30. http://www.intel.com/pressroom/archive/speeches/GEM93097.HTM).

MM—Molecular mechanics (above).

MOPAC—A software package for semi-empirical molecular orbital calculations developed by J.J.P. Stewart.

MOZYME—A modification of MOPAC that can handle very large molecules (e.g., enzymes) efficiently because it uses tricks to make the algorithm for solving for the molecular orbitals much more efficient, taking advantage of the sparsity of the Fock matrix that must be solved.

MP2 or MP4—Møller-Plesset 2nd- or 4th-order perturbation theory. Widely used approaches to include electron correlation in quantum mechanical electronic structure calculations.

MPP—Massively parallel processing.

MSI—Molecular Simulations, Inc. (above).

Multi-scale simulations—Simulations that combine approaches to allow different levels of theory with different levels of accuracy and length scales in a single simulation.

Multi-time-step molecular dynamics—Molecular dynamics where the numerical integrations are carried out using a different time step for the fast motions (e.g., bond stretching) and slow motions (e.g., van der Waals interactions).

NEMD—Non-equilibrium molecular dynamics. Methods for molecular dynamics that can simulate systems far from equilibrium; useful in studying transport properties, particularly those related to viscosity and flow.

NIH—National Institutes of Health, Bethesda, MD. See http://www.nih.gov.

NIST—National Institute of Standards and Technology, Gaithersburg, MD and Boulder, CO. A research arm of the U.S. Department of Commerce. See http://www.nist.gov; http://www.doc.gov.

NSF—National Science Foundation, Arlington, VA. U.S. government agency that provides funding for basic science and engineering research. See http://www.nsf.gov.

NMR—Nuclear magnetic resonance. An indispensable tool for structure determination of small molecules and, more and more, macromolecules, using the absorption and relaxation of nuclear spins in a permanent magnet, subjected to a radio-frequency field.

NRCC—National Resource for Computation in Chemistry. A user facility at Lawrence Berkeley Laboratory which in 1978 was one of the first large U.S. efforts to develop computational chemistry.

NWChem—A computational-chemistry software package being developed by Pacific Northwest National Laboratory and its collaborators, emphasizing efficient calculation for parallel computers. See http://www.emsl.pnl.gov:2080/docs/nwchem/.

O(N) methods—Modeling methods whose size and expense scale linearly with the number of electrons.

ONIOM methods—Methods for including different parts of the system at different levels of theory, in order to allow more atoms to be included, but the central parts to be included more accurately; developed by Morokuma and co-workers at Emory University.

ORNL—Oak Ridge National Laboratory, U.S. DOE national laboratory in Oak Ridge, TN. See http://www.ornl.gov/.

Oxford Materials Ltd. —A British software company focused on materials simulation. See http://www.oxmat.co.uk/.

Oxford Molecular Group—A software company purchased by Tektronix and most recently by Pharmacopeia; major products have included CAChe (http://www.cachesoftware.com/, now a Fujitsu product), UNICHEM, Dgauss, Topkat.

Periodic methods (atom-centered and plane-wave)—Methods to put the periodicity (e.g., in the solid state) into either molecular dynamics / Monte Carlo or quantum mechanical calculations.

Pharmacokinetics—The ADME processes of drugs, involving both reactions and mass transport.

Pharmacogenomics—Selection of a drug by knowing that some percentage of the population has a certain genomic trait.

Proteomics—The study of proteins, their separation, and their analysis.

Q-Chem—A group of *ab initio* software packages developed and marketed by Q-Chem Inc. (http://www.q-chem.com/).

QM/MM methods—Methods to combine quantum mechanical electronic structure theory for the central part of the system, using molecular mechanics for the remainder.

QM—Quantum mechanics.

QSAR/QSPR—Quantitative structure activity/property relationships. Approaches to quantify and derive a mathematical model relating molecular structural, electronic or partitioning properties to overall macroscopic chemical properties or biological activities.

Reactive Monte Carlo—A technique that permits the incorporation of chemical equilibrium into Monte Carlo simulations.

Receptors—The chemical structures on cell membranes that bind with chemical features of molecules, signal transmitters, disease organisms, and so on.

Relaxation times—The time scales of a molecular system in which processes occur.

SAC—Safety assessment candidate. A compound that is deemed worthy of preclinical safety analysis and is likely to go on to be a drug.

Scalable algorithms—Computational algorithms that can take advantage of parallel computers and, thus, can perform significantly faster on N processors than 1 processor.

Scanning tunneling microscopy (STM)—A method to derive structures of systems in the nanometer length scale.

SD—Slater determinant (defined in Eq. 2.3).

Self-consistent reaction field (SCRF) methods—Continuum methods to calculate the solvation free energies of a molecular system, often combined with quantum mechanical electronic structure calculations to derive solvation properties of the molecular, such as solvation free energies, pKa's, etc.

Semi-empirical molecular orbital methods—Methods to solve the Schroedinger equation for the electronic structure of molecules, characterized by deriving some parameters from experimental data and approximating others.

Semichem, Inc.—A software company (Shawnee, KS) specializing in semi-empirical quantum mechanics, quantitative structure/activity relationships (QSAR), and graphics for input and visualization. See http://www.semichem.com/.

Simulated annealing—Optimization methods to get to lower energy states in a system more efficiently by raising the temperature and then lowering it during the simulation. Can be used with either molecular dynamics or Monte Carlo methods.

SPARTAN—Multi-platform software package including capabilities both of molecular mechanics and of quantum-mechanical electronic structure calculations; sold by Wavefunction, Inc. (Irvine CA; http://www.wavefunction.com/).

STM—Scanning Tunneling Microscopy.

Structure-based ligand design—Approaches to design optimum or novel ligands to bind to a macromolecular target in cases where the structure of the target is known.

SYBYL—Software package of Tripos, Inc. intended for drug design and discovery, combining structural bioinformatics, chemical informatics, virtual high-throughput screening, and structure-activity studies. See http://www.tripos.com/.

Target—In drug discovery, a genetic sequence or protein that is used as a basis for intervening in a disease process.

Three-dimensional QSAR, 3DQSAR—The use of electrostatic and or steric fields around a molecule in order to derive a mathematical relationship between these fields and biological activity (see COMFA).

Tight-binding QM—Calculation of electronic contribution to interatomic forces that solves an eigenvalue problem for the matrix that represents interference between electronic orbitals rather than involve electronic wave functions explicitly.

Transition state—The structure that is the highest free energy between reactant(s) and product(s) in a chemical reaction but a minimum in all other degrees of freedom, representing the transition in identity from reactant(s) to product(s).

Tripos, Inc.—A software company (St. Louis, MO) founded in 1979 that sells software mainly for the pharmaceutical/biotech market. See http://www.tripos.com/.

UA—United-atom model for molecular simulations, in which hydrogen atoms are included lumped into a CH, CH_2, or CH_3 group (vs. EA or EH).

Van der Waals interactions or London dispersion forces—Non-bonded interactions between atoms that have an attractive component (dispersion attraction) and a repulsive component (exchange repulsion). In a Lennard-Jones 12-6 model of the overall potential, the former goes as the inverse 6^{th} power of the distance; the latter as the inverse 12^{th} power.

VASP—An *ab initio* solid-state software package (http://tph.tuwien.ac.at/~vasp/) developed at the University of Vienna.

Wavefunction, Inc.—Molecular-modeling software company (Irvine CA) founded in 1991; aims at multi-platform codes (Unix, PC, Mac) for industrial and academic markets. See http://www.wavefunction.com/.

Wien97—An *ab initio* solid-state software package developed at the University of Vienna.

XANES—X-ray absorption near edge structure. Spectral oscillations at energies just higher than those required to excite bound electrons in an absorbing atom; gives information about vacant orbitals, electronic configuration and the symmetry of the atom's site.

XPLOR—A software package developed at Harvard and Yale Universities that has become the de-facto standard for use in refining NMR and X-ray structures.

X-ray crystallography—A method to determine molecular structure of small molecule and macromolecule crystals by diffracting X-rays off the crystal. Probably the largest number of Nobel prizes have involved using this method than any other single method.

APPENDIX E. INDEX OF COMPANIES

3

3M, Inc., viii, 40, 95, 142, 159, 318, 319, 320

A

Abbott Laboratories, 337
Agouron, 52, 55, 56
Agouron Pharmaceuticals, 337
Air Liquide, viii, 40, 94, 119, 159, 180, 181, 182, 224
Air Products and Chemicals, Inc., viii, xiii, 40, 59, 95, 97, 142, 159, 180, 181, 322, 323, 324, 325, 326, 354
Akzo Nobel, 97, 124
Albemarle Chemicals, 142
Alcoa, 142
Allied Signal, 72, 109
American Cyanamid, 142
American Home Products, 337
Ames Laboratory, 154, 331
ARCO Chemicals, 142
Argonne National Laboratory, 68, 72, 154, 328
Army Research Laboratory (ARL), 153
Arqule, 56
Asahi Chemical Industry Co., Ltd., viii, x, xiii, 95, 133, 159, 261, 263, 264, 265, 267
AstraZeneca, 251, 254
Avecia, 129
Aventis, 119, 247

B

B.F. Goodrich, 142, 354
BAE (British Aerospace Engineering), 129, 206, 210
Banyu Pharmaceutical Co., Ltd., 133
BASF AG, viii, x, xi, xv, 40, 43, 44, 94, 97, 122, 159, 183, 202, 209, 210, 228, 249, 354
Bayer AG, viii, x, xi, 40, 44, 57, 94, 122, 142, 159, 188, 189, 193, 395
BG Technology Ltd., 94, 97, 159, 194
BNFL, 129, 252, 254
BOC, 97
BP, 40, 95, 129, 142, 210, 251, 252, 326, 327, 354
BP (Amoco), x, xi, 202
Bristol-Myers Squibb, 142, 337, 348
Brookhaven National Laboratory, 154

C

Calgon Inc., 143
CambridgeSoft.com, 143, 337, 338
CEA (Commissariat à l'Énergie Atomique), 120, 149
Celanese, 40
Celanese AG, 247
Center for Atomic-Scale Materials Physics (CAMP Denmark), 15, 16, 83, 90, 119, 148, 196, 198, 243, 317, 420
Centre Européen de Calcul Atomique et Moléculaire (CECAM), 117, 148, 149, 159, 199, 200, 202, 249
Centre National de la Recherche Scientifique (CNRS), 94, 120, 121, 149, 150, 199, 201, 224, 225, 236, 258, 259
Chamot Laboratories, Inc., 143
ChemInnovation Software, 331
Chevron Corporation, viii, 143, 159, 327, 339, 397
Ciba Specialty Chemicals, 129
Clorox, 382
Colgate Palmolive, viii, 129, 143, 159
Computing for Science, Ltd., 155, 420
Corning, xi, 245, 382
Corus, 129, 210
COSMOlogic, viii, xi, 122
CourtAulds plc, 129, 210

D

DaimlerChrysler, 122, 184, 246
Daresbury Laboratory, 19, 60, 70, 71, 94, 116, 129, 130, 148, 149, 150, 204, 206, 207, 208, 209, 210, 211, 212, 237, 305, 420
Degussa-Hüls AG (Degussa), x, 40, 94, 122, 213
Delphi Research Laboratories, 143
Dow-Corning Corporation, 40, 43, 143, 345, 346, 347
DSM Research, x, 40, 94, 124, 215, 216, 217, 241, 243, 354
DuPont Pharmaceuticals Company, x, 54, 56, 97, 143, 348, 349, 350, 351, 352

E

E. I. du Pont de Nemours and Company, viii, x, 40, 45, 57, 92, 95, 97, 144, 159, 210, 242, 243, 351, 352, 353, 354
Eastman Chemical Company, x, xiii, 5, 40, 97, 144, 356, 357
Eastman Kodak, 144
Eisai Co. Ltd., x, 133, 268
Elf Atochem, 202
Eli Lilly and Co., 252, 254
EniChem S.p.A., x, 94, 124, 219, 220, 399
EniTecnologie, 89, 94, 124, 222, 223
Environmental Protection Agency, 156
Exxon, 144, 354
ExxonMobil, 40, 144, 145, 327

F

FMC, 129

Ford Motor Company, viii, 1, 46, 97, 144, 159, 184, 358
Fuji Photo Film Co., Ltd., viii, 133, 159, 265, 266, 315
Fuji Xerox Co., Ltd., 133
Fujisawa Pharmaceutical Co., Ltd., 133
Fujitsu Ltd. and Fujitsu Laboratories, viii, xi, 40, 43, 60, 71, 72, 74, 129, 133, 144, 159, 210, 265, 270, 271, 272, 273, 274, 281, 289, 293, 304, 328, 349, 419, 423

G

Gaussian, Inc., viii, 46, 54, 83, 144, 159, 191, 192, 214, 216, 219, 223, 233, 237, 248, 249, 251, 256, 259, 320, 323, 331, 337, 342, 346, 353, 377, 380, 393, 398, 421
GdR 12090 project, 46, 94, 98, 120, 149, 150, 224, 225, 235, 238, 242, 244, 258
General Electric Corporation, 72, 109, 144, 184, 363, 364
General Mills, Inc., 144
General Motors, 109
Glaxo Wellcome plc, 129, 226, 227, 256, 337
GlaxoSmithKline, 129, 130, 226, 251, 256

H

Haldor Topsøe, viii, 90, 93, 94, 97, 119, 159, 196, 197, 243, 246
Halliburton Energy Services, Inc., 144
Harwell Atomic Energy Establishment, 70
Henkel, 382
Hercules Corporation, 144
Hoechst, 40, 145, 184, 202, 243, 247, 354
Hoffmann-La Roche (HLR), 56, 128, 228, 229, 230
Honeywell, 72
HRL Laboratories, 145, 365
Hypercube, Inc., 54, 378, 422

I

IBM, 44, 46, 49, 59, 60, 67, 69, 70, 71, 72, 128, 145, 153, 165, 206, 210, 359
IBM, 231
ICI, 40, 109, 129, 206, 209, 210, 250, 251, 252, 253, 420
Institut Français du Pétrole, x, 40, 46, 94, 98, 120, 148, 149, 150, 202, 224, 225, 235, 258, 259
Institute for Molecular Science, 136, 140, 151, 175, 275, 276, 278, 287, 289, 412
Institute for Solid State Physics (ISSP), 60, 70, 71, 72, 123, 136, 151, 275, 280, 281, 282, 283, 284
Intel, 71, 72
Intevep (Petróleos de Venezuela), 141

J

Janssen Pharmaceutica, 118, 191, 192, 193
Japan Polyolefins, Ltd. (Showa Denko K.K.), 95, 134, 265, 315
Japan Tobacco Inc., 134
Johnson & Johnson, 337
Joint Research Center for Atom Technology (JRCAT), 26, 40, 60, 70, 71, 95, 98, 108, 136, 151, 267, 272, 288, 289, 290, 291, 312
JSR Corporation, 134, 315

K

Kellogg, viii, 159

L

L'Oréal, 119
Lawrence Berkeley National Laboratory, 149, 154, 421, 422, 423
Lawrence Livermore National Laboratory, 63, 154, 155, 422
Lion Biosciences AG, 189
Los Alamos National Laboratory, 154, 155
Louisiana State University, 2, 59, 69, 176, 367
Lubrizol, viii, 1, 45, 46, 145, 159
Lucent Technology / Bell Laboratories, viii, 59, 65, 69, 72, 145, 159, 167, 374, 375, 376, 377

M

Marathon Oil Company, 95, 145, 378
Materials Design s.a.r.l., 119
Merck & Co., Inc., viii, xiii, 145, 159, 253, 337, 379, 395
Mexican Petroleum Corporation (PEMEX), 141
Millenium Pharmaceuticals, x, 145
Mitsubishi Chemicals, viii, x, 8, 40, 45, 55, 95, 134, 159, 265, 292, 293, 294, 295, 299, 337, 399
Mitsubishi-Tokyo Pharmaceuticals Inc., 134
Mitsui Chemicals, 95, 134, 135, 264, 265, 315, 316, 317
Mobil, 145
Mobil (ExxonMobil), 323
Molecular Simulations, Incorporated (MSI), viii, 46, 53, 54, 56, 79, 83, 89, 97, 109, 112, 129, 130, 146, 153, 159, 182, 186, 187, 188, 189, 193, 194, 206, 209, 210, 213, 214, 216, 219, 223, 224, 226, 236, 237, 244, 245, 248, 253, 254, 259, 260, 263, 266, 268, 293, 294, 309, 319, 320, 323, 324, 328, 330, 331, 337, 342, 345, 349, 350, 353, 361, 363, 364, 382, 383, 384, 393, 394, 395, 396, 408, 419, 420, 422, 423
Monsanto, 97, 146, 408
Motorola, Inc., viii, xi, 59, 69, 71, 72, 111, 145, 159, 369, 387, 388, 389

N

NASA, 66, 157
National Cancer Institute, 156
National Institute of Standards and Technology (NIST), xvi, 1, 2, 14, 59, 95, 112, 142, 152, 154, 170, 177, 331, 386, 390, 391, 392, 397, 398, 399, 419, 423

National Science Foundation (NSF), 157
Naval Research Laboratory (NRL), 67, 153
NEC Tsukuba Laboratories, xi, 26, 60, 70, 71, 72, 74, 95, 134, 135, 151, 237, 267, 275, 286, 289, 296, 297, 298, 299, 312, 314, 317
Network Science Corporation, 146
NIH Center for Molecular Modeling, 156
Nippon Steel, 109
Nippon Zeon Co., Ltd., 95, 134, 265, 315, 316
Norsk Hydro, 126, 208, 209, 210, 242, 420
Novartis Pharmaceuticals AG, viii, 128, 159, 184, 228, 229, 395

O

Oak Ridge National Laboratory, 72, 153, 154, 155, 423
Ohmine Research Group, 137, 301, 302
Owens Corning, viii, 159, 393
Oxford Materials Ltd., 129, 210, 250, 423
Oxford Molecular Group, 229, 271, 309, 349, 382, 394, 398, 419, 423

P

Pacific Northwest National Laboratory, 49, 155, 156, 329, 331, 421, 423
Pfizer, 130, 252, 254, 337
Pharmacopeia, Inc., viii, x, xvi, 56, 129, 146, 159, 167, 169, 330, 337, 382, 394, 395, 396, 422, 423
Philips, 70, 71, 72, 125, 210, 241, 242
Phillips Petroleum Company/Chevron Phillips Chemical Company, x, 95, 146, 152, 397, 399, 412
Praxair, 146, 181
Procter & Gamble, 146, 337

Q

Q-Chem, Inc., 146, 423

R

Reaction Design, Inc., 331
Rhodia (Rhône-Poulenc), 202
Rhône-Poulenc (Aventis), 40, 119
Rhône-Poulenc (Rhodia), viii, xiii, 40, 45, 97, 120, 159, 202, 241, 247, 248, 249
Rhône-Poulenc Industrialisation SA, 148
RIKEN—The Institute of Physical and Chemical Research, 26, 40, 136, 151, 206, 303, 304, 305, 306
Robert Bosch GmbH, 122
Rockwell International, 72
Rohm and Haas Company, viii, xiv, 40, 146, 159, 400, 401, 402, 412
Royal Society of Chemistry Molecular Modeling Group, 37, 129, 130, 131, 158, 250, 251, 254, 328

S

Sandia National Laboratories, 59, 69, 71, 147, 154, 156, 175, 329, 331, 387, 388, 403, 404, 405, 406
Sankyo Co., Ltd., 134
Sanofi-Synthélabo, 120
Schering-Plough, 395
Schlumberger, 130, 252
Schrödinger, Inc., 146, 323, 342, 353, 356, 422
Schroedinger, Inc., 254, 380
Semichem, Inc., 46, 398
Shell, 95, 97, 109, 125, 146, 184, 202, 206, 210, 241, 242, 243, 245, 354
Siemens, 210
Siemens AG, 122
SmithKline Beecham Pharmaceuticals, 130, 226, 256, 257, 310
SOHIO, 109
Solutia Inc., 146, 408, 409
Solvay, 94, 97, 118, 191, 193, 242
SSCI, Inc., 146
Statoil, x, 126
Steacie Institute for Molecular Science, 148
Sumitomo Chemical Co., Ltd., viii, x, 40, 95, 134, 135, 159, 265, 299, 315, 316, 317

T

Taisho Pharmaceutical Co., Ltd., 54, 133, 135, 307, 308
Takeda Chemical Industries Ltd., viii, 54, 135, 159, 309, 310
Tanabe Pharmaceutical Co., Ltd., 135
Texas Instruments, 59, 146, 369
The Dow Chemical Company, viii, x, xiii, xvi, 40, 43, 45, 46, 95, 97, 109, 143, 147, 159, 178, 341, 342, 343, 344, 345, 354, 356
Toshiba Corporation, viii, xi, 60, 71, 72, 95, 135, 159, 265, 311, 312, 313
TotalFina (TotalFinaElf), viii, x, 40, 94, 97, 120, 159, 224, 225, 242, 244, 258, 259
Toyota Central Research & Development Laboratories, Inc, 135, 314
Tripos, Inc., 52, 53, 54, 187, 189, 254, 309, 353, 363, 395, 424

U

UBE Industries, Ltd., viii, 95, 135, 159, 265, 315, 317
Unilever, viii, x, xi, 26, 40, 43, 44, 112, 129, 130, 159, 206, 210, 241, 242, 249, 252, 260, 382
Union Carbide, x, 40, 85, 87, 89, 95, 97, 147, 354, 399
Uniroyal Chemical Ltd., 140
University of Antwerp, 118, 191
University of Minnesota Supercomputing Institute, 59, 397, 398, 401, 410, 411, 412
UOP, 147

V

Vertex Pharmaceuticals, x, 52, 55, 56, 147, 167, 253, 337
VTT (Valtion Teknillinen Tutimuskeskus, 119

W

W.R. Grace, viii, 147, 159
Wavefunction, Inc., 54, 147, 216, 219, 256, 320, 323, 342, 346, 356, 363, 424, 425
Wright-Patterson AFB, Aeronautics Systems Center, 153

X

Xerox, viii, 46, 109, 153, 159
Xerox (Canada), 140

Y

Yamanouchi Pharmaceutical Co., Ltd., 135

Z

Zeneca, 395
Zeneca Specialty Chemicals, 129, 206, 210

JTEC/WTEC reports[1] are available on the Web at http://www.wtec.org, or from the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22161. Call for prices and availability (703-487-4650; FAX 703-321-8547).

JTEC Panel Report on High Temperature Superconductivity in Japan (11/89) PB90-123126

JTEC Panel Report on Space Propulsion in Japan (8/90) PB90-215732

JTEC Panel Report on Nuclear Power in Japan (10/90) PB90-215724

JTEC Panel Report on Advanced Computing in Japan (10/90) PB90-215765

JTEC Panel Report on Space Robotics in Japan (1/91) PB91-100040

JTEC Panel Report on High Definition Systems in Japan (2/91) PB91-100032

JTEC Panel Report on Advanced Composites in Japan (3/91) PB90-215740

JTEC Panel Report on Construction Technologies in Japan (6/91) PB91-100057

JTEC Panel Report on X-Ray Lithography in Japan (10/91) PB92-100205

WTEC Panel Report on European Nuclear Instrumentation and Controls (12/91) PB92-100197

JTEC Panel Report on Machine Translation in Japan (1/92) PB92-100239

JTEC Panel Report on Database Use and Technology in Japan (4/92) PB92-100221

JTEC Panel Report on Bioprocess Engineering in Japan (5/92) PB92-100213

JTEC Panel Report on Display Technologies in Japan (6/92) PB92-100247

JTEC Panel Report on Material Handling Technologies in Japan (2/93) PB93-128197

JTEC Panel Report on Separation Technology in Japan (3/93) PB93-159564

JTEC Panel Report on Knowledge-Based Systems in Japan (5/93) PB93-170124

NASA/NSF Panel Report on Satellite Communications Systems & Technology (7/93): Executive Summary (PB93-231116); Analytical Chapters (PB93-209815); Site Reports (PB94-100187)

WTEC Monograph on Instrumentation, Control & Safety Systems of Canadian Nuclear Facilities (7/93) PB93-218295

JTEC/WTEC Annual Report and Program Summary 1993/94 (3/94) PB94-155702

JTEC Panel Report on Advanced Manufacturing Technology for Polymer Composite Structures in Japan (4/94) PB94-161403

ITRI Monograph on Benchmark Technologies Abroad: Findings From 40 Assessments, 1984-94 (4/94) PB94-136637

WTEC Panel Report on Research Submersibles and Undersea Technologies (6/94) PB94-184843

JTEC Panel Report on Microelectromechanical Systems in Japan (9/94) PB95-100244

WTEC Panel Report on Display Technologies in Russia, Ukraine, and Belarus (12/94) PB95-144390

JTEC Panel Report on Electronic Manufacturing and Packaging in Japan (2/95) PB95-188116

JTEC Monograph on Biodegradable Polymers and Plastics in Japan (3/95) PB95-199071

JTEC Panel Report on Optoelectronics in Japan and the United States (2/96) PB96-152202

JTEC Panel Report on Human-Computer Interaction Technologies in Japan (3/96) PB96-157490

WTEC Panel Report on Submersibles and Marine Technologies in Russia's Far East and Siberia (8/96) PB96-199526

JTEC Panel Report on Japan's ERATO and PRESTO Basic Research Programs (9/96) PB96-199591

JTEC/WTEC Panel Report on Rapid Prototyping in Europe and Japan: Vol. I. Analytical Chapters (3/97) PB97-162564
Vol. II. Site Reports (9/96) PB96-199583

WTEC Panel Report on Advanced Casting Technologies in Japan and Europe (3/97) PB97-156160

WTEC Panel Report on Electronics Manufacturing in the Pacific Rim (5/97) PB97-167076

WTEC Panel Report on Power Applications of Superconductivity in Japan and Germany (9/97) PB98-103161

WTEC Workshop Report on R&D Status and Trends in Nanoparticles, Nanostructured Materials, and Nanodevices in the United States (1/98) PB98-117914

WTEC Panel Report on Electronic Applications of Superconductivity in Japan (7/98) PB98-150139

WTEC Monograph on Use of Composite Materials in Civil Infrastructure in Japan (8/98) PB98-158215

WTEC Panel Report on Nanostructure Science and Technology (12/98 — available from Kluwer Academic Publishers)

WTEC Panel Report on Global Satellite Communications Technology and Systems (12/98) PB99-117954

WTEC Panel Report on Digital Information Organization in Japan (2/99) PB99-128019

WTEC Panel Report on the Future of Data Storage Technologies (6/99) PB99-144214

WTEC Panel Report on Japan's Key Technology Center Program (9/99) PB99-142424

WTEC Workshop on Russian R&D Activities on Nanoparticles and Nanostructured Materials (12/99) PB99-150518

WTEC/MCC Strategic Technology Tour Report on MEMS and Microsystems in Europe (1/2000) PB2000-104945

WTEC Panel Report on Wireless Technologies and Information Networks (7/2000) PB2000-105895

WTEC Panel Report on Environmentally Benign Manufacturing (4/2001) PB2001-104339

[1] Note: see the WTEC WWW server (http://www.wtec.org) for a list of earlier JTECH reports (1984-89) available from NTIS.